IRINA-DRAGA CĂRUNTU

MORFOMETRIE COMPUTERIZATĂ MICROSCOPICĂ
ÎN HISTOLOGIE ŞI HISTOPATOLOGIE

EDITURA "GR. T. POPA" - U.M.F. IAŞI
2003

Scanned with OKEN Scanner

IRINA-DRAGA CĂRUNTU

Domnului Prof. dr. Valeriu Rusu,
cu mulțumiri pentru încrederea acordată
și sentimente de aleasă considerație,
Irina

MORFOMETRIE COMPUTERIZATĂ MICROSCOPICĂ

ÎN HISTOLOGIE ŞI HISTOPATOLOGIE

Scanned with OKEN Scanner

EDITURA "GR. T. POPA"
UNIVERSITATEA DE MEDICINĂ ŞI FARMACIE
str. Universităţii nr. 16, IAŞI

Referenţi ştiinţifici:
Prof. dr. doc. Gioconda DOBRESCU – U.M.F. "Gr. T. Popa" Iaşi
Prof. dr. Coriolan COTUŢIU – U.M.F. "Gr. T. Popa" Iaşi
Dr. Ştefan SPULBER – U.M.F. "Carol Davila" Bucureşti

Tehnoredactare: Smaranda Acatrinei, Irina-Draga Căruntu
Redactor: ing. Dana Bugeac

Descrierea CIP a Bibliotecii Naţionale a României
CĂRUNTU, IRINA-DRAGA
Morfometrie computerizată microscopică în histologie şi histopatologie / Irina-Draga Căruntu – Iaşi: Editura Gr. T. Popa, 2003
Bibliogr.
Index.
ISBN 973-7906-34-9

611.018

Scanned with OKEN Scanner

CUVÂNT ÎNAINTE

Lucrarea "Morfometrie computerizată microscopică" este dedicată unui domeniu modern de investigare în sfera histologiei şi histopatologiei, care polarizează interesul a numeroase colective din universităţi de prestigiu şi căruia îi este acordat un spaţiu de publicare tot mai larg. Acest domeniu are o istorie recentă, individualizându-se ca direcţie de cercetare în anii `90. Actuala sa dezvoltare este impulsionată atât de progresele teoretice care fundamentează "vederea" computerizată, cât şi de evoluţia tehnologiilor hardware şi software, motiv pentru care teritoriul său ştiinţific nu este complet conturat încă. Din acelaşi motiv, pe termen scurt şi mediu este uşor de întrevăzut o dinamică foarte alertă, cu o extindere amplă a perspectivelor de utilizare şi un impact cert asupra practicii diagnosticului microscopic.

Elaborată în spiritul tendinţelor curente de cercetare din acest domeniu şi reflectând nivelul prezent de cunoaştere, lucrarea are un pronunţat caracter de noutate ştiinţifică pentru literatura medicală din România. Pe lângă noutatea domeniului în sine, cartea realizează în premieră, prin concepţie şi organizare, o grefare sistematică a noţiunilor şi tehnicilor generale de procesare numerică a imaginilor în exclusivitate pe fondul imagistic specific morfologiei microscopice.

Concepută ca un suport teoretic pentru implementările tehnicilor digitale în histologie şi histopatologie, cartea reflectă punctul de vedere şi interesele ştiinţifice ale specialistului în morfologie microscopică, oferind deopotrivă clinicienilor şi cercetătorilor bazele morfometriei computerizate.

Se nasc şi la Iaşi specialişti în morfometrie !

24 februarie 2003

Acad. Laurenţiu M. Popescu

Scanned with OKEN Scanner

În amintirea bunicilor mei,

de la care am deprins profunzimea gândirii și rigoarea raționamentului matematic.

PREFAȚĂ

Cartea de față își propune să realizeze un studiu cuprinzător al tehnicilor digitale utilizate în procesarea informațiilor vizuale specifice preparatelor microscopice de histologie şi histopatologie. Această abordare asistată de calculator este în principal orientată pe metodele moderne de morfometrie, care operează la nivel de pixel şi asigură o acuratețe deosebită pentru analiza cantitativă.

Structurarea materialului în patru capitole şi două anexe a urmărit atingerea unor obiective parțiale, formulate de aşa manieră încât să asigure, pe de o parte, continuitatea expunerii şi, pe de altă parte, transmiterea unui summum de cunoştințe cu finalitate globală în înțelegerea şi manipularea complexității inerente aplicațiilor concrete. Capitolul 1 joacă un rol introductiv, concentrându-se asupra caracterizării imaginilor de morfologie microscopică şi marcând trecerea de la morfometria clasică la morfometria computerizată, ce permite creşterea vitezei şi preciziei în măsurare. Capitolul 2 este proiectat ca o cheie de boltă a întregii construcții, prezentând procesarea digitală a imaginilor şi tehnici morfometrice de înaltă precizie, adaptate la contextul morfologiei normale şi patologice. Capitolul 3 completează discutarea fundamentelor teoretice care constituie suportul morfometriei computerizate, tratând interpretarea evaluărilor cantitative prin prisma utilizării noțiunilor adecvate din teoria probabilităților şi statistică.

Pornind de la ansamblul de informații furnizat în capitolele anterioare, capitolul 4 exemplifică, prin cinci studii de caz demonstrative, proiectarea şi exploatarea modulelor software ce răspund unor probleme uzuale din practica morfometriei microscopice. Anexa I abordează pe scurt procesarea numerică a imaginilor color, fiind separată de restul textului datorită faptului că, deşi indispensabilă pentru morfologia normală şi patologică, utilizarea culorilor nu este ilustrată efectiv în lucrare (date fiind costurile foarte ridicate ale reproducerilor color). Anexa II pune la dispoziția cititorilor textele sursă şi schemele logice ale modulelor software implementate în Zeiss KS400, concepute pentru studiile de caz din capitolul 4; comentariile detaliate asupra codurilor au drept scop facilitarea înțelegerii acestora, reprezentând, totodată, modele de programare.

Natura explorărilor pe imagini microscopice de morfologie normală şi patologică diferă substanțial de problematica abordată în studiile tehnice. Deosebirile majore față de imagistica tehnică decurg din densitatea extrem de mare a elementelor prezente în scenă, precum şi din faptul că delimitarea lor este mult mai vagă, ambele deosebiri reflectându-se în dificultatea implementărilor dedicate morfologiei. Urmare a unor testări minuțioase, pe imagini din colecții personale, lucrarea este capabilă să depăşească, prin soluții originale, dificultatea menționată, fructificând o bibliografie ce include, deopotrivă, cărți de referință pentru vederea computerizată, de sorginte tehnică, şi articole de dată recentă, raportând cercetări în morfometria biomedicală.

Scanned with OKEN Scanner

Prin modul în care este concepută, cu numeroase explicații, detalieri şi exemple ilustrative, lucrarea poate fi utilizată şi drept material didactic pentru studierea tehnicilor actuale, asistate de calculator, destinate prelucrării şi analizei imaginilor biomedicale. În particular, poate oferi suport pentru cursuri postuniversitare sau la nivelul planurilor de învățământ specifice facultăților cu profil medical.

În baza elementelor definitorii ale lucrării, enunțate anterior, sfera cititorilor vizați cuprinde studenți cu aptitudini pentru activitatea de cercetare, studenți şi tineri absolvenți preocupați de aplicabilitatea informaticii medicale, medici anatomo-patologi interesați de actualizarea cunoştințelor profesionale, prin prisma valorificării tehnicii de calcul.

Cu speranța că această lucrare, grație facturii interdisciplinare a subiectului, va contribui la deschiderea unor noi perspective pentru cititorii de formație medicală, consider oportun a aprecia excelenta colaborare cu Editura "Gr. T. Popa" U.M.F. Iaşi.

Deşi ideea valorificării tehnicilor de procesare numerică a imaginilor în morfologia microscopică a constituit o preocupare personală mai veche, această idee s-a cristalizat într-un proiect de cercetare propriu-zis cu aproximativ un an în urmă. Parcurgerea distanțelor de la documentare la experimentare şi de la viziunea medicală la tehnologia programării s-a dovedit, nu rareori, plină de asperități.

Dacă azi manuscrisul vede lumina tiparului este şi pentru că, în finalizarea lui, am beneficiat de suportul unor personalități marcante din domeniul morfologiei microscopice. Astfel, adresez alese mulțumiri domnului academician Laurențiu M. Popescu, pentru disponibilitatea de "a ura drum bun" acestei cărți în lumea cititorilor, girând calitatea ei ştiințifică. De asemenea, rămân îndatorată celor trei referenți, doamna prof. dr. doc. Gioconda Dobrescu, domnul prof. dr. Coriolan Cotuțiu şi domnul dr. Ştefan Spulber, pentru timpul şi energia investite atât în parcurgerea textului, cât şi în formularea unor sugestii care au condus la creşterea comprehensibilității acestuia.

Incursiunea morfologului în universul vederii computerizate a fost susținută necondiționat de toți cei apropiați, față de care îmi exprim întreaga gratitudine.

Autoarea

Scanned with OKEN Scanner

CUPRINS

CUVÂNT ÎNAINTE I

PREFAŢĂ III

CUPRINS V

ABSTRACT AND CONTENTS IX

CAPITOLUL 1
INTRODUCERE ÎN ANALIZA CANTITATIVĂ A IMAGINILOR 1

1.1. SUBIECTIV ŞI OBIECTIV
ÎN EVALUĂRILE CANTITATIVE ALE IMAGINILOR 1

1.2. REPERE FUNDAMENTALE ÎN ANALIZA CANTITATIVĂ
A IMAGINILOR MICROSCOPICE BIOMEDICALE 4

1.2.1. ANALIZA CANTITATIVĂ CLASICĂ 4

1.2.2. ANALIZA CANTITATIVĂ MODERNĂ 5

1.3. METODE CLASICE DE MORFOMETRIE ŞI ADAPTAREA LOR
CA TEHNICI AUTOMATE DE MĂSURARE 7

1.3.1. MORFOMETRIE PRIN NUMĂRARE DE PUNCTE
(BAZATĂ PE GRILE-TEST) 7

1.3.2. EŞANTIONUL DE PROBE, SELECTAREA CÂMPURILOR,
PRECIZIA MĂSURĂTORILOR 13

Scanned with OKEN Scanner

CAPITOLUL 2
PROCESAREA DIGITALĂ A IMAGINILOR ŞI TEHNICI MORFOMETRICE DE ÎNALTĂ PRECIZIE 15

2.1. IMAGINI DIGITALE DE MICROSCOPIE OPTICĂ 16
2.1.1. DIGITIZAREA IMAGINILOR 16
2.1.2. PROBLEMATICA REZOLUŢIEI IMAGINII DIGITALE 17
2.1.3. STOCAREA IMAGINILOR DIGITALE 18
2.1.4. CRITERII ÎN SELECTAREA CÂMPURILOR PENTRU IMAGINI DIGITIZATE DE ÎNALTĂ CALITATE 19

2.2. CONCEPTE ŞI INSTRUMENTE 20
2.2.1. ETAPELE PROCESĂRII NUMERICE A IMAGINII 21
2.2.2. HISTOGRAMA ŞI UTILIZAREA EI ÎN CARACTERIZAREA IMAGINILOR DIGITALE 21
2.2.3. VECINĂTĂŢILE UNUI PIXEL 24
2.2.4. ELEMENTE STRUCTURALE SAU MĂŞTI 25
2.2.5. REGIUNI ÎNTR-O IMAGINE DIGITALĂ 26

2.3. PROCESAREA PRELIMINARĂ – TEHNICI DE ÎMBUNĂTĂŢIRE A IMAGINILOR DIGITALE 27
2.3.1. AJUSTAREA LUMINOZITĂŢII 27
2.3.2. MANIPULAREA CONTRASTULUI 29
2.3.3. ELIMINAREA ZGOMOTELOR ŞI UMBRELOR 38
2.3.4. DETECŢIA MUCHIILOR ŞI IDENTIFICAREA CONTURURILOR. SEPARAREA REGIUNILOR 42

2.4. SEGMENTAREA IMAGINILOR DIGITALE 46
2.4.1. BINARIZAREA DE TIP PRAG 47
2.4.2. BINARIZAREA PRIN CREŞTEREA REGIUNILOR 51

2.5. OPERAŢII ASUPRA IMAGINILOR BINARE 53
2.5.1. MORFOLOGIE BINARĂ 53
2.5.2. OPERAŢII BOOLEENE 59

2.6. MĂSURARE / MĂSURĂTORI 62
2.6.1. DESCRIPTORI GEOMETRICI STANDARD 63
2.6.2. UNITĂŢI DE MĂSURĂ ŞI CONVERSII 65
2.6.3. DESCRIPTORI BIFAZICI SPECIFICI MORFOMETRIEI CELULARE 67
2.6.4. UTILIZAREA PARAMETRILOR GEOMETRICI BIDIMENSIONALI ÎN INTERPRETĂRI STEREOLOGICE 67

Scanned with OKEN Scanner

CAPITOLUL 3
INTERPRETAREA MATEMATICĂ A REZULTATELOR MĂSURĂTORILOR 71

3.1. NOŢIUNI DESPRE VARIABILELE ALEATOARE 71
3.1.1. NOŢIUNEA DE PROBABILITATE 72
3.1.2. VARIABILE ALEATOARE CU UN NUMĂR FINIT DE VALORI 72
3.1.3. VARIABILE ALEATOARE CU UN NUMĂR INFINIT DE VALORI 74
3.1.4. CORELAŢIA A DOUĂ VARIABILE ALEATOARE 80

3.2. ELEMENTE DE STATISTICĂ 81
3.2.1. POPULAŢIE STATISTICĂ, EŞANTION SAU SELECŢIE 82
3.2.2. SERII STATISTICE ŞI VARIABILE ALEATOARE EMPIRICE 82
3.2.3. REPREZENTAREA GRAFICĂ A SERIILOR STATISTICE 84
3.2.4. ELEMENTE CARACTERISTICE ALE SERIILOR STATISTICE 88
3.2.5. DISTRIBUŢII TEORETICE ŞI DISTRIBUŢII EXPERIMENTALE 92
3.2.6. DEPENDENŢĂ STOHASTICĂ – METODA CORELAŢIEI ŞI A REGRESIEI 94

CAPITOLUL 4
STUDII DE CAZ SEMNIFICATIVE 99

4.1. SCURTĂ PREZENTARE A UTILIZĂRII MEDIULUI ZEISS KS400 ÎN DOMENIUL BIOMEDICAL 100

4.2. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU STRUCTURI CELULARE. ILUSTRARE LA NIVELUL PULPEI DENTARE 103
4.2.1. DESCRIEREA PROCEDURII 103
4.2.2. INTERPRETAREA REZULTATELOR MORFOMETRICE 107

4.3. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU STRUCTURI TUBULARE ÎN SECŢIUNE TRANSVERSALĂ. ILUSTRARE LA NIVELUL DENTINEI 110
4.3.1. DESCRIEREA PROCEDURII 110
4.3.2. INTERPRETAREA REZULTATELOR MORFOMETRICE 114

4.4. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU STRUCTURI TUBULARE ÎN SECŢIUNE LONGITUDINALĂ. ILUSTRARE LA NIVELUL DENTINEI 116
4.4.1. DESCRIEREA PROCEDURII 116
4.4.2. INTERPRETAREA REZULTATELOR MORFOMETRICE 120

Scanned with OKEN Scanner

4.5. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU REGIUNI DE INTERES ÎN INTERIORUL UNUI CÂMP. ILUSTRARE PRINTR-UN ADENOCARCINOM ENDOMETRIAL 123

4.5.1. DESCRIEREA PROCEDURII 123

4.5.2. INTERPRETAREA REZULTATELOR MORFOMETRICE 127

4.6. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU ELEMENTE DE INTERES DIN INTERIORUL UNEI REGIUNI. ILUSTRARE LA NIVELUL CORPUSCULULUI RENAL 128

4.6.1. DESCRIEREA PROCEDURII 128

4.6.2. INTERPRETAREA REZULTATELOR MORFOMETRICE 133

ANEXA I
PARTICULARITĂŢI ÎN PROCESAREA IMAGINILOR COLOR 135

ANEXA II
TEXTE SURSĂ ŞI SCHEME LOGICE 139

ANEXA II.1 (ODO_P1, ODO_P2) 140

ANEXA II.2 (TUBDEN) 146

ANEXA II.3 (DENTIN, DENT) 149

ANEXA II.4 (ADK) 159

ANEXA II.5 (COREN) 163

BIBLIOGRAFIE 169

INDEX 173

Scanned with OKEN Scanner

Scanned with OKEN Scanner

COMPUTERIZED MICROSCOPIC MORPHOMETRY IN HISTOLOGY AND HISTOPATHOLOGY

— ABSTRACT AND CONTENTS —

The book aims to provide a comprehensive study of the digital techniques, used in processing visual information originated from histology and histopathology specimens. This computer-assisted approach is mainly oriented towards the modern methods of morphometry, which operate at a pixel level and ensure a maximum accuracy in the quantitative analysis. In the following, we give a brief overview of the book, in order to synthetically present our scientific endeavor. In Chapter 1 (**Introduction to the quantitative analysis of images**), the first part refers to the role played by the quantitative analysis in removing the subjectivity factor from the interpretation of images. Later on, the state of art in the quantitative analysis of morphologic images is shortly addressed. The last section details a series of basic methods belonging to the classic morphometry, which initially were meant as manual procedures (requesting the direct intervention of the human observer) and, in time, became automated (by their adaptation to the computational resources). Chapter 2 (**Digital image processing and high precision morphometry techniques**) is the center of gravity of the whole work, and it has an adequate size in the overall structure. This is why the exposition borrows key-elements and concepts from the computer vision texts, but the illustrations through examples and the discussions are exclusively placed in the histologic and histopathologic area, by using images from personal collections. In Chapter 3 (**Mathematical interpretation of the measurement results**), we discuss those basic notions from the probability theory and statistics, mandatory for a deep understanding of the measurement results, performed on light microscopy images. To provide a truly efficient material for the readers, this book contains concrete information regarding the exploitation of a software environment, namely Zeiss KS400. Thus, Chapters 2 and 3 include numerous examples with a strongly didactic character, intended to illustrate, step by step, the basic theoretical notions of the digital image processing and the statistical interpretation of the results. On the other hand, Chapter 4 (**Illustrative case studies**) offers more elaborated examples, constructed as case studies, which respond to complex problems with a certain degree of generality, frequently encountered in the current practice of the microscopic morphometry. These case studies provide enough details to permit the reproduction of all the quantitative analysis experiments, either under Zeiss KS400 or any other dedicated environment, with similar functionality.

FOREWORD I

PREFACE III

CONTENTS (IN ROMANIAN) V

ABSTRACT AND CONTENTS (IN ENGLISH) IX

CHAPTER 1
INTRODUCTION TO THE QUANTITATIVE ANALYSIS OF IMAGES 1

1.1. SUBJECTIVE AND OBJECTIVE IN THE QUANTITATIVE ASSESMENTS OF IMAGES 1

1.2. FUNDAMENTALS OF THE QUANTITATIVE ANALYSIS OF THE MICROSCOPIC BIOMEDICAL IMAGES 4

1.2.1. CLASSIC QUANTITATIVE ANALYSIS 4

1.2.2. MODERN QUANTITATIVE ANALYSIS 5

1.3. CLASSIC MORPHOMETRY METHODS AND THEIR ADAPTATION AS AUTOMATIC MEASUREMENT TECHNIQUES 7

1.3.1. POINT COUNTING MORPHOMETRY (BASED ON TEST GRIDS) 7

1.3.2. SAMPLING, FIELD SELECTION, MEASUREMENT PRECISION 13

CHAPTER 2
DIGITAL IMAGE PROCESSING AND HIGH PRECISION MORPHOMETRY TECHNIQUES 15

2.1. DIGITAL IMAGES OF LIGHT MICROSCOPY 16

2.1.1. IMAGE DIGITIZING 16

2.1.2. ON THE RESOLUTION OF DIGITAL IMAGES 17

2.1.3. STORAGE OF DIGITAL IMAGES 18

2.1.4. CRITERIA IN FIELD SELECTION FOR HIGH QUALITY DIGITIZED IMAGES 19

2.2. CONCEPTS AND INSTRUMENTS 20

2.2.1. STAGES IN DIGITAL IMAGE PROCESSING 21

2.2.2. THE HISTOGRAM AND ITS USAGE IN THE CHARACTERIZATION OF DIGITAL IMAGES 21

2.2.3. PIXEL NEIGHBOURHOODS 24

2.2.4. STRUCTURAL ELEMENTS OR MASKS 25

2.2.5. REGIONS IN A DIGITAL IMAGE 26

Scanned with OKEN Scanner

2.3. PRELIMINARY PROCESSING – ENHANCEMENT TECHNIQUES FOR DIGITAL IMAGES 27
2.3.1. BRIGHTNESS ADJUSMENT 27
2.3.2. CONTRAST MANIPULATION 29
2.3.3. NOISE AND SHADOW REMOVAL 38
2.3.4. EDGE DETECTION AND CONTOUR IDENTIFICATION. REGION SELECTION 42
2.4. DIGITAL IMAGE SEGMENTATION 46
2.4.1. THRESHOLD SEGMENTATION 47
2.4.2. REGION GROWING SEGMENTATION 51
2.5. BINARY IMAGE OPERATIONS 53
2.5.1. BINARY MORPHOLOGY 53
2.5.2. BOOLEAN OPERATIONS 59
2.6. MEASURING / MEASUREMENTS 62
2.6.1. STANDARD GEOMETRIC PARAMETERS 63
2.6.2. MEASUREMENT UNITS AND CONVERSIONS 65
2.6.3. TWO-PHASE PARAMETERS IN CELL MORPHOMETRY 67
2.6.4. USING BIDIMENSIONAL GEOMETRIC PARAMETERS IN STEREOLOGIC INTERPRETATIONS 67

CHAPTER 3
MATHEMATICAL INTERPRETATION OF THE MEASUREMENT RESULTS 71

3.1. BASIC INFORMATION ON RANDOM VARIABLES 71
3.1.1. THE CONCEPT OF PROBABILITY 72
3.1.2. RANDOM VARIABLES WITH A FINITE NUMBER OF VALUES 72
3.1.3. RANDOM VARIABLES WITH AN INFINITE NUMBER OF VALUES 74
3.1.4. THE CORRELATION OF TWO RANDOM VARIABLES 80
3.2. ELEMENTS OF STATISTICS 81
3.2.1. STATISTICAL POPULATION, SAMPLE OR SELECTION 82
3.2.2. STATISTICAL SERIES AND EMPIRICAL RANDOM VARIABLES 82
3.2.3. GRAPHICAL PLOTS ASSOCIATED WITH STATISTICAL SERIES 84
3.2.4. CHARACTERISTICS OF STATISTICAL SERIES 88
3.2.5. THEORETICAL AND EXPERIMENTAL DISTRIBUTIONS 92
3.2.6. STOCHASTIC DEPENDENCE – THE CORRELATION AND REGRESSION METHODS 94

Scanned with OKEN Scanner

CHAPTER 4
ILLUSTRATIVE CASE STUDIES 99

4.1. BRIEF OVERVIEW OF ZEISS KS400 USAGE IN BIOMEDICAL DOMAIN 100

4.2. QUANTITATIVE MORPHOLOGICAL ANALYSIS FOR CELLULAR STRUCTURES. ILLUSTRATION ON THE DENTAL PULP 103

4.2.1. PROCEDURE DESCRIPTION 103

4.2.2. INTERPRETATION OF THE MORPHOMETRIC RESULTS 107

4.3. QUANTITATIVE MORPHOLOGICAL ANALYSIS FOR TUBULAR STRUCTURES IN CROSS SECTION. ILLUSTRATION ON DENTIN 110

4.3.1. PROCEDURE DESCRIPTION 110

4.3.2. INTERPRETATION OF THE MORPHOMETRIC RESULTS 114

4.4. QUANTITATIVE MORPHOLOGICAL ANALYSIS FOR TUBULAR STRUCTURES IN LONGITUDINAL SECTION. ILLUSTRATION ON DENTIN 116

4.4.1. PROCEDURE DESCRIPTION 116

4.4.2. INTERPRETATION OF THE MORPHOMETRIC RESULTS 120

4.5. QUANTITATIVE MORPHOLOGICAL ANALYSIS FOR REGIONS INSIDE A FIELD. ILLUSTRATION ON AN ENDOMETRIAL ADENOCARCINOMA 123

4.5.1. PROCEDURE DESCRIPTION 123

4.5.2. INTERPRETATION OF THE MORPHOMETRIC RESULTS 127

4.6. QUANTITATIVE MORPHOLOGICAL ANALYSIS FOR ELEMENTS INSIDE A REGION. ILLUSTRATION ON RENAL CORPUSCLE 128

4.6.1. PROCEDURE DESCRIPTION 128

4.6.2. INTERPRETATION OF THE MORPHOMETRIC RESULTS 133

APPENDIX I
PARTICULARITIES IN COLOR IMAGE PROCESSING 135

APPENDIX II
SOURCE CODES AND FLOWCHARTS 139

REFERENCES 169

INDEX 173

Scanned with OKEN Scanner

Scanned with OKEN Scanner

CAPITOLUL 1

INTRODUCERE ÎN ANALIZA CANTITATIVĂ A IMAGINILOR

Precizia ridicată pe care o asigură în prezent *morfometria biomedicală computerizată* este o consecință a dezvoltării analizei cantitative a imaginilor (înglobând principii și instrumente de măsurare), în paralel cu ansamblul de tehnici digitale ce permit o prelucrare rafinată a imaginilor. Drept urmare, un prim obiectiv al acestui capitol îl constituie sublinierea rolului jucat de analiza cantitativă în eliminarea subiectivismului din interpretarea unor imagini de interes general. Ulterior, discuția referitoare la analiza cantitativă se axează numai pe imagini de morfologie microscopică, realizându-se o succintă trecere în revistă a stadiului actual din sfera histologiei și histopatologiei. Ultima parte a capitolului curent detaliază o serie de proceduri de bază ale morfometriei clasice, inițial practicate manual (cu participarea nemijlocită a observatorului uman) și, în timp, automatizate (prin adaptarea la facilitățile sistemelor de calcul). Considerăm oportună prezentarea unor atare proceduri în acest context cu caracter introductiv, spre a crea o viziune completă asupra evoluției metodelor de măsurare, cu toate că abordarea modernă a analizei cantitative nu face propriu-zis apel la ele (după cum va reieși din capitolul următor).

1.1. SUBIECTIV ŞI OBIECTIV ÎN EVALUĂRILE CANTITATIVE ALE IMAGINILOR

Modalitatea prin care inteligența umană recunoaşte o imagine are la bază o asociere cu un model (în sensul de imagine-model), ca urmare a vizualizării anterioare repetate a unor imagini aparţinând aceloraşi clase de obiecte/structuri. Sunt construite astfel **modele mentale** ale obiectelor, iar informaţia corespunzătoare imaginii percepute şi recunoscute, împreună cu alte informaţii despre obiect, face posibilă încadrarea într-un anumit tip de clasificare.

Modelele mentale permit operarea cu un anumit grad de obiectivitate, existând însă întotdeauna şi elemente subiective. Câteva exemple despre elementele subiective – pe care în general nu le percepem conştient – sunt edificatoare (Turek, 2000):

- suntem sensibili la diferenţele de contrast, motiv pentru care supraestimăm dimensiunea unui obiect dacă prezintă un contrast înalt, comparativ cu un contrast scăzut;
- suntem sensibili la transformările de perspectivă şi profunzime;
- suntem sensibili la orientarea surselor de lumină, având o mai bună percepţie în cazul în care lumina vine de sus;
- de multe ori, involuntar, completăm imaginea cu informaţii despre care gândim că ar trebui să existe.

În acest context, prezentăm în continuare o serie de exemple, încadrabile în clasa iluziilor optice (Turek, 2000), care demonstrează subiectivitatea percepţiei ochiului uman în raport cu evaluările cantitative.

Exemplul 1.1.1

În 1861, Ewald Hering publică desenul din fig. 1.1.1, suscitând un interes deosebit din partea astronomilor vremii, care, confruntaţi cu imaginea respectivă, şi-au dat seama că observaţiile bazate pe percepţiile vizuale nu sunt întotdeauna reale. Imaginea creează impresia de profunzime (datorită dispoziţiei radiale a liniilor, precum şi faptului că acestea sunt punctate), iar cele două linii orizontale par a fi uşor curbate în zona centrală. Prin îndepărtarea fundalului format din liniile radiale punctate, iluzia dispare. Percepţia corectă a imaginii trebuie realizată însă pe imaginea completă (păstrând fundalul radial); şi în această situaţie, se poate ajunge la concluzia paralelismului prin măsurarea, într-un număr de puncte, a distanţei dintre cele două linii, distanţele rezultând riguros egale.

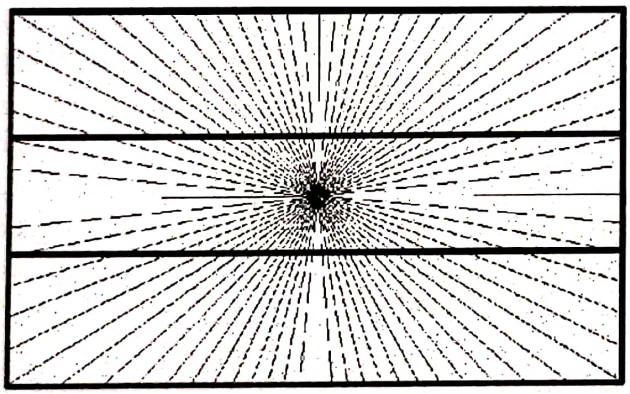


Fig. 1.1.1. Iluzia Hering (adaptare după Turek, 2000).

Exemplul 1.1.2

Imaginea din fig. 1.1.2 prezintă o altă iluzie bazată pe profunzime, sesizată în 1889 de Franz Müller-Lyon. Subiectiv, lungimea identică a celor două linii verticale este percepută diferit de un observator uman, datorită inversării modalităţii de dispunere a liniilor oblice de la cele două extremităţi. Iluzia este legată de capacitatea creierului de a procesa informaţia de perspectivă, pentru a determina profunzimea şi distanţa, transformând, aparent, o imagine bidimensională într-una tridimensională. Dacă se îndepărtează din imaginea iniţială liniile oblice, iluzia dispare. Obiectiv însă, pe imaginea completă, măsurarea celor două linii verticale demonstrează faptul că lungimile lor sunt identice.

Scanned with OKEN Scanner

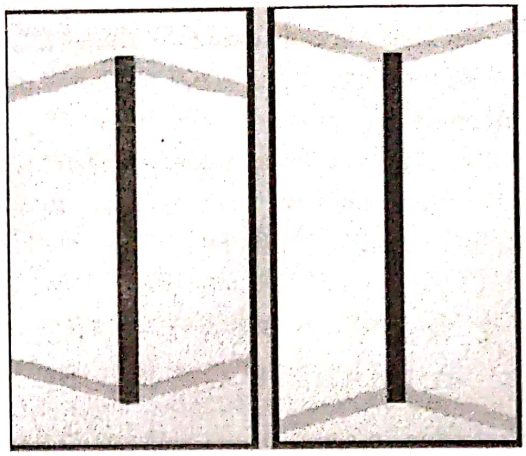


Fig. 1.1.2. Iluzia Müller-Lyon (adaptare după Turek, 2000).

Exemplul 1.1.3

Imaginea din fig. 1.1.3(a) prezintă o altă iluzie optică, sesizată în 1860 de Johann Poggendorff. Cele două segmente ale liniei diagonale par a fi uşor deplasate unul faţă de altul. Prin îndepărtarea dreptunghiului localizat în centrul imaginii şi suprapus peste linia diagonală, iluzia dispare şi se observă că, în realitate, linia este continuă. Operând pe imaginea completă şi trasând (eventual automat) o dreaptă care se suprapune peste segmentul din porţiunea inferioară a imaginii, vom constata că această dreaptă se va suprapune şi peste segmentul din porţiunea superioară a imaginii – după cum indică linia albă din fig. 1.1.3(b).

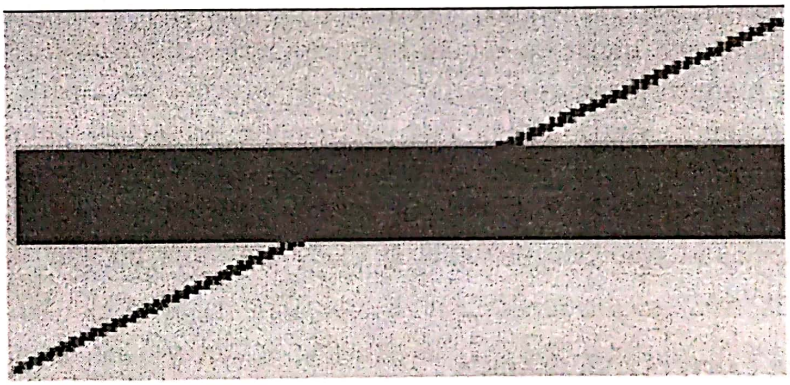


Fig. 1.1.3(a). Iluzia Poggendorff (adaptare după Turek, 2000).

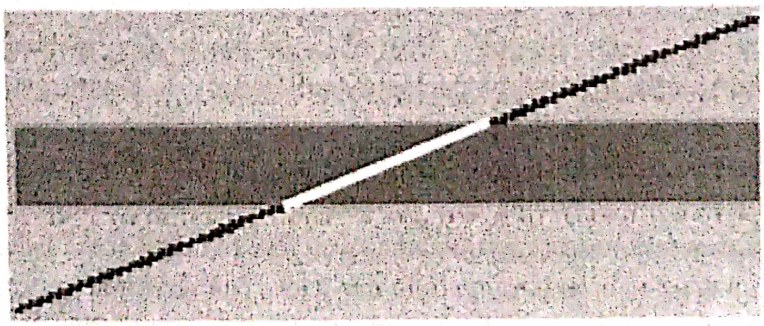


Fig. 1.1.3(b). Refacerea continuităţii liniei.

Scanned with OKEN Scanner

Toate aceste exemple ilustrează existența unui grad de **subiectivitate** în percepția vizuală, datorat factorului uman. Gradul de subiectivitate poate fi contracarat prin **obiectivitatea** procesului de măsurare, chiar dacă acesta este manual. Mai mult, în cazul analizei digitale de imagine, evaluarea cantitativă a informațiilor se realizează de către un echipament (sistem tehnic), care funcționează cu premise de totală obiectivitate (fără implicarea factorului uman).

1.2. REPERE FUNDAMENTALE ÎN ANALIZA CANTITATIVĂ A IMAGINILOR MICROSCOPICE BIOMEDICALE

Practica histologiei şi histopatologiei implică două condiții esențiale: o bază teoretică solidă, formată printr-o documentare științifică riguroasă, şi o experiență practică vastă, acumulată printr-un proces instructiv, în care capacitatea de vizualizare ocupă un rol decisiv. *Morfologia microscopică este predominant o știință a vederii colorate*, medicii patologi utilizând extensiv vederea în scop diagnostic şi depinzând de ochii lor, responsabili de peste 90% din rezultatele activității (Alvira, 2000). Din punct de vedere fiziologic însă, vederea umană este predominant calitativă şi comparativă, mai degrabă decât cantitativă. Oricât de bogată ar fi experiența unui morfolog, există limite în diferențierea caracteristicilor de culoare sau luminozitate în imagini. În baza acestor afirmații, se consideră că procesul analitic primar în morfologia microscopică este în principal calitativ şi comparativ.

Totuşi, trebuie luată în considerare şi caracterizarea **cantitativă** a imaginilor. Extragerea şi evaluarea cantitativă a trăsăturilor unor structuri dintr-o imagine pot fi aplicate atât în activitatea curentă de diagnostic, cât şi în activitatea de cercetare științifică, domeniile de interes fiind reprezentate de citologie, histologie şi histopatologie. În cazul unei investigații citologice, structurile sunt tratate ca entități bidimensionale. În histologie şi histopatologie, pe lângă tratarea bidimensională, structurile trebuie interpretate şi ca părți ale unei realități tridimensionale (Olinici, Vaida, 1997). Prin informațiile de factură numerică pe care le furnizează, rezultând pe seama evaluării unor criterii proprii morfologiei microscopice, metodele cantitative sunt deosebit de utile în formularea diagnosticului anatomo-patologic pozitiv şi diferențial în patologia tumorală şi în cuantificarea parametrilor care definesc leziunile border-line.

1.2.1. ANALIZA CANTITATIVĂ CLASICĂ

Morfometria, ca metodă de analiză cantitativă în medicină, şi în principal în sfera morfologiei, datează din a doua jumătate a secolului XX, tehnicile utilizate fiind limitate ca şi performanțe.

Morfometria a fost introdusă de Weibel în (Weibel, 1969), care o definește drept descrierea cantitativă a unei structuri. Utilizarea în practică a morfometriei presupune procedee de cuantificare a trăsăturilor (sau caracteristicilor) de natură geometrică relevante pentru respectiva structură, în ipoteza când organizarea tridimensională a acesteia este cunoscută (în sensul de corect înțeleasă). În baza literaturii de specialitate, se poate afirma că aplicarea morfometriei în medicină are în vedere două metode principale: planimetria şi stereologia. **Planimetria** (morfometria plană) se referă la măsurarea trăsăturilor geometrice corespunzătoare imaginilor bidimensionale (chiar dacă, fapt subînțeles, structurile respective nu întotdeauna pot fi considerate bidimensionale). Conform definiției din (Weibel, 1979), **stereologia** pune la dispoziție procedee ce permit obținerea unor informații tridimensionale

Scanned with OKEN Scanner

despre structura studiată, utilizând imagini bidimensionale; cu alte cuvinte, rezultatele măsurătorilor efectuate pe imagini bidimensionale sunt extrapolate, în sensul caracterizării tridimensionale a structurilor (în cazul satisfacerii anumitor ipoteze de lucru) (Taşcă, 1976, Gundersen, Osterby, 1981, Gundersen et al, 1988, Weibel, 1989, Howard, Reed, 1998, Campbell, 2002 a, e).

În numeroase situații practice, analiza morfometrică de tip planimetrie furnizează specialistului anatomo-patolog suficiente date pentru aprecierea cantitativă a preparatului microscopic, ceea ce face ca abordarea stereologică să nu fie întotdeauna necesară.

Tehnica de morfometrie în care măsurătorile sunt proiectate şi realizate efectiv de către un observator uman, care, în final, furnizează rezultatele determinărilor numerice, poate fi denumită **morfometrie clasică**; ea se bazează pe aplicarea de **grile-test** pentru numărare (principiile fundamentale vor fi expuse în secțiunea 1.3).

1.2.2. ANALIZA CANTITATIVĂ MODERNĂ

Informațiile furnizate prin măsurătorile morfometrice sunt cu atât mai valoroase, cu cât măsurătorile respective sunt mai rafinate. O atare rafinare, care conduce la o investigare performantă a caracteristicilor cantitative celulare şi tisulare, este realizată la ora actuală utilizând facilitățile oferite de dezvoltarea tehnicii de calcul. Acest tip de analiză cantitativă este denumit **morfometrie asistată de calculator** şi constituie un instrument de lucru la care se face apel în mod frecvent.

În general, utilizarea calculatoarelor personale a devenit o practică curentă de anatomie patologică în ultimii ani, ca urmare a reducerii substanțiale a costurilor echipamentelor şi programelor aferente. Prin aceasta, resursele hardware (echipamente) şi software (programe) necesare producerii şi prelucrării imaginilor digitale sau numerice (specifice modului de operare numeric – conform termenului anglo-saxon "digit", însemnând "cifră") au dobândit o largă accesibilitate. Atributul "digital" sau "numeric" asigură delimitarea fermă față de noțiunea de **imagine "analogică"**, explicații suplimentare urmând a fi date în capitolul 2.

În esență, o imagine analogică nu poate face obiectul memorării (stocării) pe un sistem de calcul, datorită caracterului de continuitate spațială. Merită de menționat faptul că imaginile-sursă utilizate în morfologie sunt de natură eminamente analogică (de exemplu, imaginile din câmpul microscopic al preparatelor histologice sau histopatologice, imaginile din fotografii pe suport de hârtie etc.).

În raport cu **precizia măsurătorilor**, se diferențiază **două clase** de metode în morfometria asistată de calculator:

- În cazul primei clase, se aplică modul de operare din morfometria clasică, numai că grilele-test de numărare se suprapun automat peste imaginile din calculator, iar cuantificarea este realizată tot automat, de către calculator; observatorul uman intervenind în delimitarea structurilor de interes (Russ, Dehoff, 2001, Russ, 2002, Slomianka, 2002, Systems Planning and Analysis, Inc., 2002). Precizia măsurătorilor este dată de precizia grilei.
- În cazul celei de a doua clase, se aplică o metodologie complexă, denumită generic "**procesare (prelucrare) numerică de imagine**" sau "**procesare (prelucrare) de imagine asistată de calculator**". Metodologia presupune o succesiune de etape (detaliate în capitolul 2), în cadrul căreia analiza cantitativă reprezintă punctul final.

Scanned with OKEN Scanner

Precizia măsurătorilor este incontestabil mai bună decât cea obținută prin utilizarea grilelor, fiind impusă de modul de operare **la nivel de pixel**, specific procesării numerice a imaginilor. (Pixelul constituie unitatea indivizibilă din alcătuirea unei imagini digitale – pentru definirea exactă, se va vedea secțiunea 2.1 a capitolului 2.)

În contextul celor sus-menționate, în subsecțiunea 1.2.1 și în subsecțiunea curentă, trebuie subliniat faptul că termenul de "analiză cantitativă" se suprapune peste (este echivalat cu) cel de "morfometrie".

În cazul analizei cantitative de imagine asistată de calculator, complexitatea implementărilor poate varia de la nivelul unor programe simple, al căror control necesită intervenția utilizatorului, până la aplicații ce includ elemente de inteligență artificială, de tipul **sistemelor expert** (Tosi, Cottier, 1989, Collan et al., 1992).

Pe lângă efectuarea automată a măsurătorilor (ca urmare a procesării imaginilor), un sistem expert dedicat morfologiei microscopice este capabil să interpreteze rezultatele respectivelor măsurători, îndeplinind astfel sarcini care, de obicei, necesită intervenția inteligenței umane (Baak, 1988).

Facilitățile de această factură oferite de sistemul expert sunt asigurate de posibilitatea interogării automate a unor baze de date și a unor baze de cunoștințe, construite anume (în faza de proiectare a sistemului expert) pentru a păstra informații numerice și principii calitative ce constituie esența experienței practice a morfologului în formularea diagnosticului. În consecință, aplicația software, în ansamblul ei, acționează asemănător unui "expert" în probleme de patologie, ceea ce, din punct de vedere etimologic, a condus la apariția denumirii de "sistem expert", conferită acestor tipuri de aplicații.

Astfel, pentru un același tip de analiză referitor la imagini histologice sau citologice similare, un sistem expert utilizează un același algoritm de prelucrare și respectă aceleași reguli decizionale (Connely, 1990). De obicei, criteriile decizionale (parte componentă a bazei/bazelor de cunoștințe) apelează la tehnici de recunoaștere a formelor (*eng.* pattern recognition) specifice domeniului larg al inteligenței artificiale, în sensul că trăsăturile imaginii curente (ce face obiectul analizei) sunt confruntate cu trăsăturile unor imagini (sau modele) de referință (memorate în baza/bazele de date cu care a fost dotat sistemul expert).

În urma unei atare confruntări, orice sistem expert fiabil (robust) asociază specificitatea imaginii curente cu un model de referință (eventual mai multe modele de referință), iar în cazul când premisele asocierii nu sunt îndeplinite în limita unor erori admisibile, situația respectivă este comunicată operatorului uman. Prin această strategie de funcționare operatorul uman are posibilitatea de a introduce elemente noi în bazele de cunoștințe și/sau de date, ce suplimentează conținutul informațional inițial cu care au fost prevăzute în faza de proiectare, conferind performanțe decizionale ce acoperă mai larg procesul de recunoaștere.

După cum am precizat deja, principiile moderne de analiză cantitativă, operând la nivel de **pixel,** se încadrează în cea de a doua clasă de metode identificate mai sus. Complexitatea fundamentelor teoretice pe care se construiește această clasă face obiectul unei abordări detaliate în capitolul viitor, care, totodată, este ilustrat prin numeroase exemple concrete. Nu putem însă trece cu vederea faptul că, sub raportul principiilor, prima clasă de metode se bazează pe *exploatarea morfometriei clasice pentru elaborarea unor procedee de măsurare automatizată*. De aceea, considerăm oportună o scurtă descriere a acestora, pe parcursul secțiunii următoare.

Scanned with OKEN Scanner

1.3. METODE CLASICE DE MORFOMETRIE ŞI ADAPTAREA LOR CA TEHNICI AUTOMATE DE MĂSURARE

Prezentăm în continuare câteva principii din **morfometria clasică** (Davey, 2002) ce stau la baza efectuării manuale de măsurători, ţinând cont, totodată, că parte dintre ele (cum ar fi, de exemplu, măsurătorile bazate pe grile) pot fi utilizate şi pentru dezvoltarea unor tehnici de analiză cantitativă automată (asistată de calculator).

Când analizează un preparat microscopic, morfologul trebuie să aibă întotdeauna în vedere faptul că toate structurile sunt tridimensionale şi că o aceeaşi imagine bidimensională particulară poate să rezulte din mai multe tipuri de structuri tridimensionale. De exemplu, o structură bidimensională circulară poate să rezulte din secţionarea unei structuri tridimensionale fie cilindrice, fie sferice, fie elipsoidale, fie neregulate. Pentru a cuantifica parametrii structurii tridimensionale este necesară identificarea/recunoaşterea exactă, astfel încât să se poată aplica formulele de calcul adecvate şi, respectiv, să se poată obţine informaţii pertinente în urma măsurătorilor.

În abordarea unei imagini, un prim pas este reprezentat de **definirea structurilor de interes**, după care urmează **stabilirea măsurătorilor** pe care dorim să le realizăm. De exemplu, referindu-ne la măsurători de arii, dacă luăm în considerare cele două structuri bidimensionale din fig. 1.3.1, atunci există mai multe posibilităţi de a formula obiectivele concrete ale măsurătorilor, cum ar fi:

- să determinăm separat ariile corespunzătoare fiecărui contur;
- să determinăm aria totală corespunzătoare celor două contururi;
- să determinăm sub formă procentuală raportul dintre aria totală corespunzătoare celor două contururi şi aria întregii secţiuni.

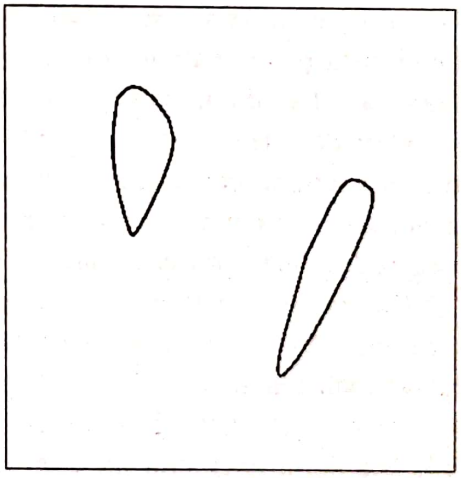


Fig. 1.3.1. Demarcarea prin contururi a două structuri de interes, într-o imagine (adaptare după Davey, 2002).

1.3.1. MORFOMETRIE PRIN NUMĂRARE DE PUNCTE (BAZATĂ PE GRILE-TEST)

Metoda de morfometrie clasică care, indiscutabil, se bucură de cea mai largă răspândire, se bazează pe **numărarea de puncte**, în scopul evaluării diverselor caracteristici ale structurilor de interes.

Din acest motiv, considerăm oportună detalierea unor principii de măsurare ce utilizează numărarea de puncte (Gundersen et al., 1981, Howard, Reed, 1998, Davey, 2002).

Scanned with OKEN Scanner

- **CALCULUL ARIILOR**

Exemplul de mai jos porneşte de la structurile de interes din fig. 1.3.1, la care se adaugă o grilă-test în pătrate, ca în fig. 1.3.2. Practic, o astfel de grilă poate fi realizată pe un material transparent, ce se suprapune peste imagine, dând posibilitatea unei raportări concomitente atât la imagine, cât şi la reperul din grilă.

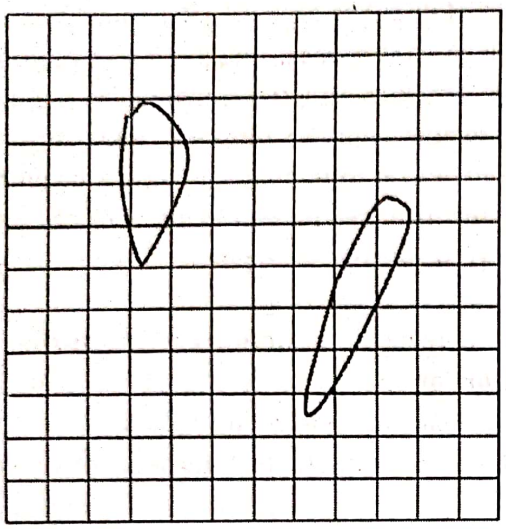


Fig. 1.3.2. Grilă de măsurare cu pătrate suprapusă peste structurile de interes (adaptare după Davey, 2002).

Astfel, problema determinării **ariei** fiecărui contur revine la a suma ariile tuturor acelor pătrate care sunt localizate predominant în interiorul contururilor. Această modalitate de măsurare implică o aproximare, care, însă, este considerată acceptabilă. Analizând structura conturată în stânga, putem aprecia că numărul de pătrate conținute în interiorul ei este 4, deşi partea de sus şi, respectiv, cea de jos a acestor 4 pătrate nu se află complet în contur. Există şi alte pătrate adiacente care conțin porțiuni ale conturului, astfel încât, prin neglijarea acestora din urmă în procesul de numărare, erorile pot să se compenseze.

O modalitate de a reduce erorile ce provin din arbitrarietatea numărării pătratelor parțial incluse în interiorul conturului o constituie operarea cu punctele din centrele de simetrie ale pătratelor grilei, aşa cum se observă în fig. 1.3.3. Astfel, se va înlocui numărarea de pătrate cu numărarea de puncte, procedeu care a condus şi la denumirea principiului metodei.

Ca urmare, vor face obiectul numărării toate punctele situate în interiorul conturului care delimitează o structură de interes. Aria structurii investigate va fi aproximată prin valoarea produsului dintre numărul de puncte şi aria unui pătrat.

Reducerea erorilor decurge din faptul că, pe de o parte, numărul pătratelor localizate strict în interiorul unei structuri este egal cu numărul de puncte din centrul acestor pătrate, şi, pe de altă parte, se elimină aprecierile legate de numărarea pătratelor "de frontieră", localizate parțial în interiorul structurii şi parțial în afara ei.

În aplicarea practică a metodei, se consideră o probabilitate redusă de apariție a situației echivoce, când există puncte situate chiar pe linia de contur. Grafic vorbind, dacă se acceptă că punctul are o anumită dimensiune, iar linia de contur are o anumită grosime, atunci acuratețea desenului constituie o premisă pentru evitarea situațiilor echivoce. Tot din rațiuni grafice, este recomandabil ca marcarea punctelor să se realizeze cu simbolul +, ceea ce, în numărarea manuală, garantează o mai bună identificare.

Scanned with OKEN Scanner

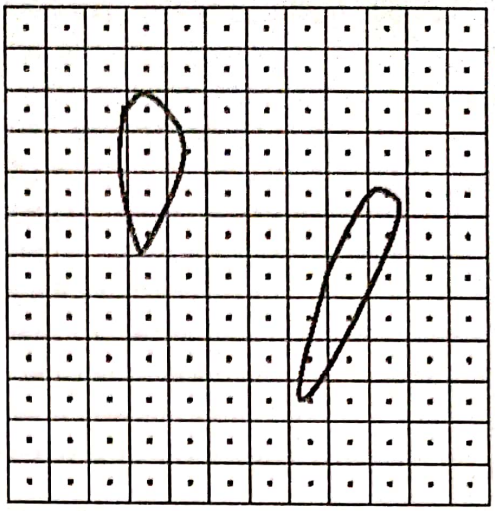


Fig. 1.3.3. Marcare cu puncte a centrelor de simetrie ale pătratelor grilei de numărare (adaptare după Davey, 2002).

În baza celor sus menționate, numărarea de puncte este preferabilă în raport cu numărarea de pătrate, ceea ce face ca în structura grilei să putem renunța la pătrate. Grila de numărare devine astfel formată numai din puncte marcate cu simbolul ".", ca în fig. 1.3.4, sau din puncte marcate cu simbolul "+", ca în fig. 1.3.5.

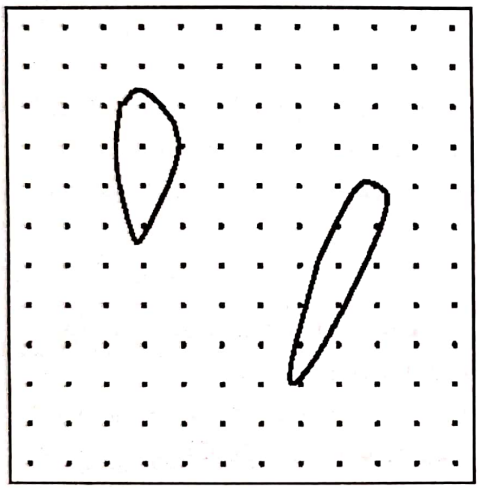


Fig. 1.3.4. Grilă de numărare realizată numai cu puncte (adaptare după Davey, 2002).

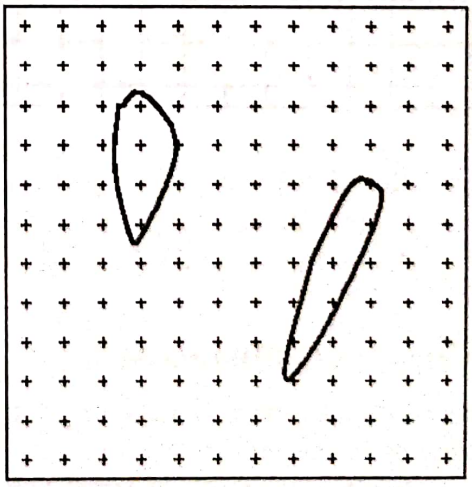


Fig. 1.3.5. Grilă de numărare realizată numai cu simboluri "+" (adaptare după Davey, 2002).

Scanned with OKEN Scanner

Tot din raţiuni legate de reducerea erorilor, grila formată din pătrate, considerată iniţial în fig. 1.3.2, poate fi utilizată prin prisma identificării vârfurilor pătratelor, ca în fig. 1.3.6. Aproximările realizate în acest mod vizează de asemenea pătratele "de frontieră", ca şi în cazul centrelor de simetrie, cu deosebirea că o caracterizare prin patru puncte (respectiv vârfurile) a unui atare pătrat devine mai rafinată decât o caracterizare printr-un singur punct (respectiv centrul de simetrie).

În aceste condiţii, pentru a estima aria structurilor este necesar să numărăm acele vârfuri ale pătratelor grilei care sunt situate în interiorul contururilor ce delimitează structurile respective, ca în fig. 1.3.6. În fig. 1.3.6 se observă 4 puncte de tip vârf pentru conturul din stânga şi 6 pentru cel din dreapta, adică în total 10 puncte de tip vârf. Calculul ariei delimitate de cele două contururi se formulează eficient ca raportare la numărul total de 144 de puncte de tip vârf de pătrat existente pe grila suprapusă peste imagine. Astfel, notând:

P (numărul punctelor de tip vârf de pătrat care aparţin structurilor de interes) = 10,

P_t (numărul total al punctelor de tip vârf de pătrat de pe grilă) = 144,

se obţine exprimarea procentuală:

$$P/P_t = 7\%.$$

Cunoscând aria totală a grilei (A_t) în unităţi fizice, se determină aria corespunzătoare celor două contururi (A) în unităţi fizice:

$$A = 0.07\, A_t.$$

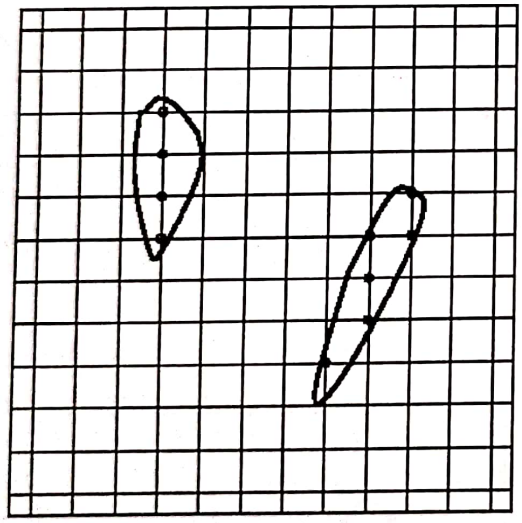


Fig. 1.3.6. Marcarea punctelor de tip vârf situate în interiorul contururilor într-o grilă de numărare realizată din pătrate (adaptare după Davey, 2002).

Observaţie. Dacă secţiunea particulară considerată este relevantă pentru structura tridimensională, atunci este posibilă extrapolarea de la raportul ariilor, la raportul volumelor (V/V_t), valoarea raportului fiind aceeaşi. Această afirmaţie se bazează pe principiile **stereologiei**.

- **CALCULUL PERIMETRELOR**

Tehnicile de numărare a punctelor pot fi utilizate pentru a determina **perimetre** (şi, implicit, rapoarte de forma perimetru/arie).

Următorul exemplu se referă la o serie de măsurători pentru un cerc, pe care le putem realiza utilizând o grilă de tipul celor prezentate anterior.

Scanned with OKEN Scanner

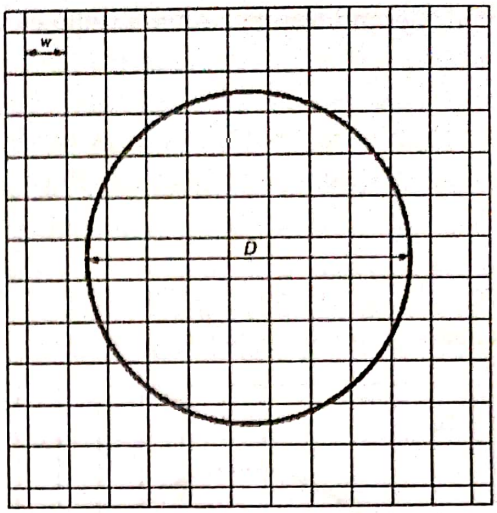


Fig. 1.3.7. Grilă de numărare formată din pătrate, suprapusă peste un cerc (adaptare după Davey, 2002).

Pornind de la desenul din fig. 1.3.7, putem calcula într-o manieră aproximativă aria, diametrul şi perimetrul cercului, după cum urmează:

- Fie P numărul total de puncte de tip centre de simetrie ale pătratelor grilei, care se află în interiorul cercului. Considerând latura pătratelor grilei de valoare w, aria cercului (calculată aproximativ) este dată de formula:

$$A = Pw^2.$$

- Fie P_D numărul de puncte de tip centre de simetrie ale pătratelor grilei situate de-a lungul diametrului orizontal al cercului. Cu aceeaşi semnificaţie pentru w, diametrul cercului (calculat aproximativ) este dat de formula:

$$D = P_D w.$$

Odată ce a fost determinat diametrul, circumferinţa cercului, care reprezintă un perimetru (calculată aproximativ), este dată de formula:

$$C = \pi D.$$

În cele ce urmează, vom furniza încă o formulă pentru calculul aproximativ al circumferinţei cercului, care însă nu foloseşte numărarea punctelor de tip centre de simetrie ale pătratelor grilei. Pentru aceasta, să considerăm cercul din fig. 1.3.8, în care, prin simbolul grafic "o", au fost puse în evidenţă punctele de intersecţie dintre cerc şi liniile grilei. Să notăm cu n_i numărul total al acestor puncte de intersecţie. Valoarea n_i poate fi calculată în baza raţionamentului de mai jos.

Un număr de D/w linii verticale ale grilei intersectează cercul, semnificaţia notaţiilor D şi w fiind cea introdusă anterior. Fiecare din aceste linii verticale are două puncte de intersecţie cu cercul (neglijând situaţiile limită ale liniilor tangente la cerc). Aşadar, liniile verticale ale grilei intersectează cercul în $2D/w$ puncte. În mod cu totul analog, se ajunge la concluzia că şi liniile orizontale ale grilei intersectează cercul tot în $2D/w$ puncte. Cu alte cuvinte, numărul total al punctelor de intersecţie dintre liniile grilei şi cerc este dat de formula:

$$n_i = 4D/w,$$

ceea ce este echivalent cu a scrie:

$$D = n_i w/4.$$

Scanned with OKEN Scanner

Totodată, din formula circumferinței cercului, diametrul D poate fi exprimat drept:

$$D = C/\pi.$$

Prin eliminarea lui D între cele două relații precedente, se obține o modalitate de calcul aproximativ a circumferinței cercului:

$$C = \pi n_i w/4,$$

care utilizează în exclusivitate informațiile puse la dispoziție de grilă (adică w și n_i) și în care nu apare, explicit, dependența de diametrul D.

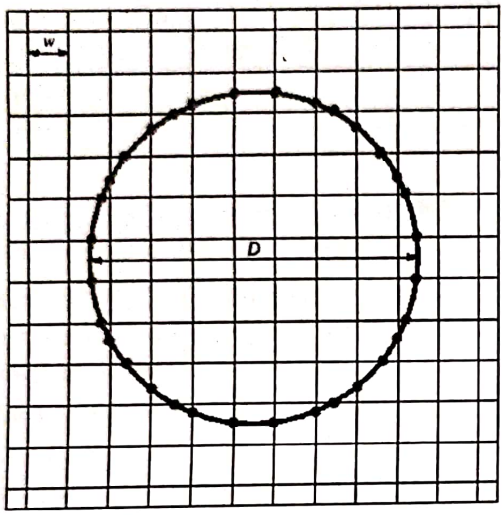


Fig. 1.3.8. Marcarea (cu simbolul grafic "o") punctelor de intersecție dintre cerc și liniile grilei de numărare (adaptare după Davey, 2002).

Comentariile privind formula anterioară sunt deosebit de importante, deoarece această formulă se poate aplica pentru contururi generate cu ajutorul unor arce de cerc, cum ar fi situația conturului prezentat în fig. 1.3.9. Arcele de cerc ce definesc conturul pot fi privite ca provenind din trei cercuri complete, ceea ce înseamnă că perimetrul corespunzător imaginii din fig. 1.3.9 este egal cu suma circumferințelor celor trei cercuri.

Constatările de mai sus, referitoare la calculul perimetrului cu ajutorul numărului de puncte de intersecție dintre liniile grilei și contur, pot fi generalizate pentru orice contur în condiții de izotropie, adică atunci când nu există o orientare preferențială a conturului.

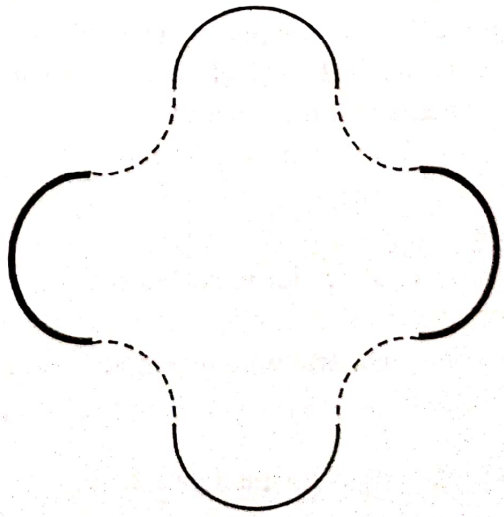


Fig. 1.3.9. Contur generat cu ajutorul unor arce de cerc de aceeași rază, al cărui perimetru este egal cu de trei ori circumferința cercului de rază dată (un cerc conturat cu linie continuă subțire, un cerc conturat cu linie continuă groasă și un cerc conturat cu linie punctată) (adaptare după Davey, 2002).

Scanned with OKEN Scanner

Observație. Prin metodele de calcul aproximativ sus menționate, ne aflăm practic și în posesia unor modalități de exprimare a rapoartelor de tipul perimetru/arie care fac apel numai la informațiile furnizate de grila de numărare. Evaluarea unor astfel de rapoarte se dovedește utilă și în cazul structurilor tridimensionale regulate, situație când se poate considera că raportul suprafață/volum este numeric egal cu raportul perimetru/arie corespunzător unei secțiuni plane. O atare afirmație se bazează pe principiile **stereologiei**. De exemplu, în cazul unei structuri tridimensionale de tip cilindric, pentru care se consideră o secțiune plană perpendiculară pe generatoarea suprafeței cilindrice, această secțiune evidențiază un cerc, ale cărui arie și perimetru se pot calcula aproximativ în conformitate cu cele discutate anterior. Astfel, putem scrie egalitatea:

$$suprafață/volum = perimetru/arie = (\pi n_i w/4)/(Pw^2) = \pi n_i/(4Pw).$$

O aplicație concretă în morfologie se referă la celulele cu structură cilindrică (fibre nervoase, celule musculare striate), sau pentru componente intracelulare (tubuli). Se poate arăta că formula obținută pentru raportul suprafață/volum rămâne valabilă în cazul mai general al structurilor tridimensionale regulate, cu contururi izotrope (nu neapărat cercuri) în secțiune transversală. În schimb, dacă structura circulară provine dintr-o structură sferică, calculul aceluiași raport va fi mult mai dificil, deoarece mărimea secțiunii circulare depinde de poziționarea secțiunii în raport cu centrul sferei.

1.3.2. EȘANTIONUL DE PROBE, SELECTAREA CÂMPURILOR, PRECIZIA MĂSURĂTORILOR

Deoarece morfometria prin numărare de puncte este inerent o tehnică de aproximație, măsurătorile pot fi limitate la un **lot** (**eșantion**) considerat **reprezentativ**. În termeni practici, se poate spune că un eșantion luat în studiu este reprezentativ atunci când rezultatele obținute prin medierea pe respectivul eșantion nu se modifică în cazul adăugării de probe (adică în cazul lărgirii eșantionului). Există totuși o serie de probleme de ordin practic cu care se confruntă patologul, legate de modul în care se selectează câmpurile semnificative ce vor fi măsurate și de precizia măsurătorilor (Howard, Reed, 1998, Campbell, 2002 a, e, Slomianka, 2002).

Selectarea câmpurilor se realizează în conformitate cu experiența observatorului și în raport cu obiectivele urmărite (activitate de diagnostic, cu evaluarea unor factori de prognostic, sau activitate de cercetare).

Precizia rezultatelor obținute printr-o tehnică de numărare depinde de mărimea grilei de numărare, în sensul că spații mai mici ale grilei asigură o precizie mai ridicată. La modul ideal, o grilă cu spații infinit de mici ar permite măsurarea exactă a ariilor, dar o atare abordare ar necesita un timp infinit de realizare a măsurătorilor.

Principiile unui compromis la care trebuie făcut apel în practică, între precizie și timpul de numărare, pot fi stabilite în urma câtorva experimente concrete. Astfel, conform recomandărilor din literatura de specialitate, dacă în câmp apar mai multe structuri similare, grila de numărare trebuie să fie suficient de mică pentru a furniza informații relevante la nivelul fiecărei structuri. De aceea, spațierea grilei trebuie aleasă în funcție de dimensiunile structurilor pentru care se efectuează măsurătorile: pentru structuri de dimensiuni mari sunt necesare grile cu spațiere largă, iar pentru structuri de dimensiuni mici – grile cu spațiere fină.

Scanned with OKEN Scanner

Este evident faptul că adaptarea tehnicilor de morfometrie clasică ca metode de analiză cantitativă asistată de calculator face ca numărarea de puncte pe o grilă test să nu mai constituie o problemă, din punctul de vedere al consumului de timp. Drept consecință, se poate alege o spațiere oricât de mică a grilei-test. Principial vorbind, dimensiunea pătratelor din grila-test ar putea fi redusă până la dimensiunea unității de reprezentare a imaginilor digitale, anume pixelul; frecvent, imaginile digitale sunt constituite din 512x512 = 262144 pixeli.

Efectuarea de măsurători la nivel de pixel exploatează în totalitate rezoluția performantă pe care o pot avea imaginile digitale, asigurând o precizie net superioară măsurătorilor realizate cu grile-test uzuale, pentru care numărul punctelor de măsurare este doar de ordinul sutelor. Astfel, *operarea la nivel de pixel stă la baza procedeelor moderne de analiză cantitativă asistată de calculator a imaginilor* și este implementată în toate mediile software de ultimă generație. Totodată, pentru a beneficia de înalta precizie asigurată măsurătorilor, aceste medii software pun la dispoziție și o serie de instrumente vizând pregătirea adecvată a imaginilor, premergătoare măsurătorilor.

Scanned with OKEN Scanner

Scanned with OKEN Scanner

CAPITOLUL 2

PROCESAREA DIGITALĂ A IMAGINILOR ŞI TEHNICI MORFOMETRICE DE ÎNALTĂ PRECIZIE

În mediile software moderne, ce permit efectuarea unor măsurători automate de înaltă precizie asupra anumitor caracteristici ale structurilor de interes dintr-o imagine microscopică, acestea nu se pot realiza direct pe imaginea inițială. Drept urmare, ea trebuie supusă unor transformări care să facă posibilă îndepărtarea informațiilor redundante în raport cu obiectivul măsurătorilor, fără a altera caracteristicile în cauză. Fiecare din aceste transformări constituie rezultatul aplicării unor operații asupra unei imagini de lucru curente. Succesiunea acestor operații, pornind de la imaginea inițială (sursă) și încheindu-se cu imaginea finală ce permite extragerea informațiilor numerice, este definită sub numele global de **procesare (prelucrare) digitală**. Derularea corectă a procesării digitale este crucială pentru validitatea rezultatelor obținute prin măsurarea caracteristicilor. În momentul când se efectuează măsurătorile, nu există nici un instrument extern de control care să garanteze că ceea ce se măsoară sunt chiar trăsăturile din imaginea inițială de la care s-a plecat. De aceea, procesarea numerică a imaginii trebuie gestionată cu maximum de responsabilitate de către utilizator, ceea ce impune o *cunoaştere în detaliu a tehnicilor specifice prelucrării*, concomitent cu o *pregătire adecvată în domeniul morfologiei microscopice*. Astfel, capitolul de față trebuie privit ca centrul de greutate al întregii lucrări, motiv pentru care i s-a acordat un spațiu adecvat în economia de ansamblu. Într-o atare structură, tratarea problematicii împrumută din stilul textelor tehnice despre vederea computerizată (*eng*. computer vision), dar, totodată, ilustrările prin exemple și discuții sunt în exclusivitate orientate în aria histologiei și histopatologiei, fiind aplicate pe imagini din colecții personale (Căruntu, 2000 a, b, Căruntu 2001 a, b).

În ciuda facilităților puse la dispoziție de platformele software moderne, dezvoltarea unei aplicații de analiză cantitativă automată, oricât de simplă, este imposibilă în lipsa unei familiarizări cu reperele standard (**principii** și **metode**) ale prelucrării numerice a imaginilor. Drept consecință, cele șase secțiuni detaliază, pentru utilizatorul din domeniul biomedical, elementele fundamentale privind *crearea și operarea cu imagini digitale*, vizând, ca finalitate, *extragerea de trăsături cantitative*. Detalierea pornește de la prezentarea

şi caracterizarea imaginilor digitale, urmate de definirea unor concepte şi instrumente esenţiale în manipularea imaginilor digitale. Ulterior, sunt aprofundate procedurile de îmbunătăţire a calităţii imaginilor digitale, operaţia de segmentare şi operaţiile cu imagini binare. Secţiunea finală este dedicată în totalitate tehnicilor actuale de morfometrie, ce garantează acurateţea măsurătorilor la nivel de pixel, însemnând o rafinare substanţială a preciziei, în raport cu morfometria bazată pe grile de tip Weibel.

În scopul înţelegerii fără dificultăţi a materialului, au fost evitate o serie de formulări matematice expresive, dar cu un grad înalt de complexitate, preferând substituirea lor prin comentarii şi aplicaţii realizate în mediile Zeiss KS400 şi MATLAB.

Atât pentru tratarea teoretică, cât şi pentru înţelegerea modalităţilor de aplicare practică, se poate face apel la numeroase surse bibliografice, dintre care amintim doar pe cele cotate ca texte de referinţă în literatura domeniului de procesare de imagine asistată de calculator (Jain, 1989, Russ, 1998, Alvira, 2000, Seul, O'Gorman, Sammon, 2000, Pratt, 2001, Shapiro, Stockman, 2001, Campbell, 2002 b, c, d, Gonzales, Woods, 2002, Russ, 2002).

2.1. IMAGINI DIGITALE DE MICROSCOPIE OPTICĂ

În procesarea de imagine asistată de calculator se operează cu **imagini digitale** (numerice), obţinute din **imaginile analogice** sursă (originale) printr-un procedeu denumit **digitizare**. Imaginile digitale au o organizare adecvată operaţiilor ce pot fi efectuate pe un calculator, precum şi modului de stocare a informaţiei specific unui astfel de echipament. Această organizare vizează, pe de o parte, rezoluţia imaginii (în sensul de claritate şi de surprindere a detaliilor spaţiale – eventual raportate la percepţia umană), şi, pe de altă parte, gama de culori sau nuanţe de gri (în sensul de claritate şi de surprindere a detaliilor coloristice – eventual raportate la percepţia umană). Câmpul sursă explorat (de natură eminamente analogică) nu poate fi selectat la întâmplare, el trebuind să satisfacă o serie de condiţii pentru ca, prin digitizare, să se ajungă la o imagine digitală de o bună calitate (prin raportare la procesarea automată ulterioară sau, eventual, la percepţia umană).

2.1.1. DIGITIZAREA IMAGINILOR

O **imagine digitală** este constituită dintr-un număr finit de elemente picturale denumite **pixeli** (*eng.* pixels), cu interdistanţa 0, organizaţi într-o manieră matriceală, similar cu pătratele unei grile de tipul celor prezentate în capitolul 1, subsecţiunea 1.3.1. Fiecărui pixel îi este asociată o anumită nuanţă de gri sau o anumită culoare din spectrul color. Astfel de asocieri se fac pe baza unor convenţii de factură numerică, specifice fiecărui echipament în parte. Pixelii, în individualitatea lor, nu conţin informaţii cu semnificaţie geometrică.

Fără a intra în detalii tehnice, menţionăm că **digitizarea** unei imagini analogice este rezultatul unei operaţii de conversie a semnalelor (informaţiilor) analogice, în semnale (informaţii) digitale. Astfel, conversia poate fi explicată pornind de la acoperirea imaginii analogice (sursă) de către o matrice (grilă) cu acelaşi număr de elemente (de obicei pătrate) şi aceeaşi organizare ca şi a pixelilor din imaginea digitală ce urmează a fi obţinută. **Conversia analog-digitală** se realizează la nivelul corespondenţei biunivoce: element din imaginea analogică – pixel din imaginea digitală. Numărul elementelor în care este astfel descompusă imaginea analogică este suficient de mare încât să se poată considera că fiecare din aceste elemente este caracterizat printr-o singură nuanţă de gri (respectiv culoare), care este memorată

Scanned with OKEN Scanner

în pixelul aferent din imaginea digitală, utilizând un anumit număr de biţi (de la termenul anglo-saxon **bit**, prescurtare de la "binary digit", însemnând cifră binară). Frecvent, numărul de biţi întrebuinţat în acest scop este 8, ceea ce permite utilizarea a $256 = 2^8$ niveluri/nuanţe de gri, respectiv de culoare (vezi Anexa I). Preferarea acestui număr de 8 biţi se datorează arhitecturilor hardware obişnuite, care pot adresa direct fiecare ansamblu de 8 biţi, denumit **octet** sau **byte**. Altfel spus, fiecărui pixel dintr-o **imagine cu nuanţe de gri** îi corespunde un octet.

Un caz particular al conversiei analog-digitale este obţinerea **imaginilor binare** sau alb-negru, când fiecărui pixel îi pot fi atribuite numai două valori, 0 sau 1, cu care se codifică albul şi, respectiv, negrul. Aşadar, reţinerea informaţiei corespunzătoare unui singur pixel necesită doar un singur bit. De obicei, imaginile alb-negru nu reprezintă rezultatul direct al unei conversii analog-digitale, în sensul că nu se justifică o digitizare cu numai două niveluri de gri (adică alb şi negru). Imaginile alb-negru sunt obţinute, în marea majoritate a situaţiilor, ca urmare a unor procedee de prelucrare a imaginilor digitale cu nuanţe de gri sau colorate (asupra cărora vom reveni ulterior cu detalii).

2.1.2. PROBLEMATICA REZOLUŢIEI IMAGINII DIGITALE

Rezoluţia imaginii este capacitatea de a distinge detaliile din respectiva imagine. În general, cu cât imaginea este mai clară, cu atât rezoluţia percepută este mai bună.

Pentru o tratare completă a problematicii rezoluţiei imaginilor digitale, trebuie avut în vedere faptul menţionat anterior, anume că imaginile digitale provin din imagini sursă de natură analogică, în urma unui proces de digitizare. Digitizarea realizează o discretizare spaţială a imaginii sursă, furnizând o organizare matriceală, care are pixelul drept element de bază.

Rezoluţia unei imagini digitale se exprimă în **număr de pixeli / unitate metrică** – cum ar fi măsurarea rezoluţiei în număr de pixeli / inch (abreviat ppi din *eng.* pixels per inch), care este utilizată curent datorită influenţei tehnologiei americane ce operează cu unitatea de tip inch (1 inch = 2.54 cm). De exemplu, o imagine cu o rezoluţie de 100 ppi conţine 10000 pixeli într-un pătrat cu latura de 1 inch (100 pixeli pe orizontală x 100 pixeli pe verticală = 10000). Dacă imaginea digitală considerată, ca mai sus, cu rezoluţie de 100 ppi, este un dreptunghi cu laturile de 3 şi, respectiv, 2 inch, atunci numărul total de pixeli corespunzător întregii imagini este 60000 = (3x100)x(2x100).

Pornind de la o aceeaşi imagine sursă (de natură analogică), imaginea digitizată va avea o rezoluţie cu atât mai bună, cu cât numărul de elemente din matricea analogică (egal cu numărul de pixeli din matricea digitală) este mai mare. De exemplu, o imagine analogică cu dimensiuni de 3 x 2 inch, digitizată cu o rezoluţie de 100 ppi, va conduce la o imagine digitală conţinând 60000 pixeli (ca în cazul analizat mai sus), în timp ce digitizarea aceleiaşi imagini cu o rezoluţie de 300 ppi va conduce la o imagine digitală conţinând 540000 = (3x300)x(2x300) pixeli.

Acest exemplu sugerează faptul că, din punct de vedere practic, digitizarea unei imagini analogice trebuie să asigure un echilibru între două cerinţe contradictorii: rezoluţia cât mai ridicată, pe de o parte, şi volumul de memorie ocupat de imaginea digitală, pe de altă parte.

Imaginile digitale utilizate de diverse programe ce pot fi executate pe calculatoarele personale se caracterizează printr-un număr de pixeli corespunzător axei orizontale şi, respectiv, un număr de pixeli corespunzător axei verticale a imaginii. Valorile curente ale celor două numere sunt 640 pixeli şi, respectiv, 480 de pixeli, dar marea majoritate a programelor permit modificarea lor şi operarea cu imagini digitale caracterizate de noile

Scanned with OKEN Scanner

valori. Este recomandabil ca alegerea celor două numere să nu fie arbitrară, căutând să se asigure o aceeaşi rezoluţie pentru axa orizontală şi axa verticală. Pornind de la o imagine digitală dată, reducerea numărului de pixeli poate conduce la o pierdere de informaţii, deoarece un pixel din noua imagine va substitui un grup de pixeli din imaginea iniţială, iar nivelul de gri (culoarea) pentru acest pixel va rezulta dintr-un proces de mediere pe grupul de pixeli aferent. În schimb, procedeul invers, de creştere a numărului de pixeli, nu poate asigura o îmbunătăţire a informaţiei din noua imagine, întrucât fiecărui pixel din imaginea iniţială i se va asocia un grup de pixeli, toţi cu caracteristici identice cu cele ale pixelului de origine. Sintetizând, prin scăderea numărului de pixeli corespunzător celor două axe, rezoluţia imaginii digitale (în raport cu imaginea sursă) va scădea, concomitent cu reducerea conţinutului informaţional, în timp ce prin creşterea numărului de pixeli rezoluţia va creşte, fără a putea îmbunătăţi conţinutul informaţional (nu pot fi aduse detalii care nu au existat în imaginea digitală pentru care s-a operat creşterea numărului de pixeli).

Vizualizarea unei imagini digitale stocate în memoria unui calculator personal, pe unul din dispozitivele standard cu care este echipat acesta (monitor, imprimantă), se realizează în conformitate cu **rezoluţia dispozitivului**, în sensul că are loc o modificare de la rezoluţia imaginii memorate, la rezoluţia imaginii ce poate fi vizualizată (afişată sau tipărită). Frecvent, se constată că unele detalii ce nu apar în imaginea afişată pe ecran sunt vizibile pe materialul tipărit, explicaţia constând în aceea că detaliile respective există în imaginea digitală memorată, iar rezoluţia imprimantei este mai bună decât cea a ecranului. De fapt, prin modificarea adecvată a rezoluţiei afişării pe monitor, atunci când este posibilă, s-ar putea vizualiza toate detaliile imaginii originale. Mai mult, o creştere exagerată a rezoluţiei de afişare pe monitor conduce, în vizualizare, la apariţia unor contururi "colţuroase" în locul unor contururi netede, datorită faptului că, pentru structura vizualizată, pe ecran se alocă mai mulţi pixeli decât în imaginea digitală memorată.

2.1.3. STOCAREA IMAGINILOR DIGITALE

Într-un echipament digital, cum este cazul unui calculator personal, o imagine digitală este **memorată** ca o succesiune de informaţii ce definesc poziţia şi nivelul de gri (în imaginile cu nuanţe de gri), respectiv, culoarea (în imaginile colorate), pentru fiecare pixel în parte. Devine astfel evident că necesarul de memorie cerut de stocarea unei imagini digitale este cu atât mai mare, cu cât numărul total de pixeli corespunzător imaginii este mai mare. Totodată, din subsecţiunea anterioară rezultă că numărul de pixeli asociaţi unei imagini digitale este direct dependent de rezoluţia şi de dimensiunile fizice ale acesteia.

O imagine digitală trebuie să memoreze mii până la milioane de pixeli, fiecărui pixel corespunzându-i un punct de alb sau negru (pentru imagini binare), un nivel de gri (pentru imagini cu nuanţe de gri) sau o culoare (pentru imagini colorate). Dacă imaginea este alb-negru, atunci este nevoie numai de un bit (o cifră binară) pe pixel, deoarece fiecare pixel poate să aibă numai o valoare din două (de exemplu 0 pentru negru sau 1 pentru alb), motiv pentru care imaginea se mai numeşte şi binară. Cu cât numărul de niveluri de gri al unei imagini monocrome (respectiv numărul de culori al unei imagini color) este mai mare, cu atât trebuie să asociem mai mulţi biţi fiecărui pixel din imagine.

În prezent, după cum s-a menţionat şi în subsecţiunea 2.1.1, cel mai frecvent se folosesc 8 biţi pe pixel, ceea ce înseamnă că se pot reprezenta 256 ($=2^8$) de nuanţe de gri distincte. O imagine ce utilizează 16 biţi pe pixel poate conţine $2^{16} = 65536$ nuanţe de gri distincte.

Scanned with OKEN Scanner

Pentru colorit de calitate profesională, denumit şi culoare adevărată (*eng.* true color), sunt utilizaţi 24 de biţi pe pixel, rezultând 2^{24} = 16777216 culori posibile, din combinaţia a trei culori primare, fiecăreia fiindu-i alocat un număr de 2^8 biţi (vezi Anexa I). De exemplu, pentru un ecran cu rezoluţie de 640x480 pixeli, utilizarea a 24 de biţi de culoare pentru fiecare pixel înseamnă un necesar de memorie de 640 x 480 x 24 = 7372800 biţi, sau echivalent, 921600 octeţi sau bytes (1 byte = 8 biţi).

Organizarea unei imagini digitale în maniera descrisă mai sus constituie un standard denumit **bitmap**, utilizabil pe toate platformele software existente în prezent. Datorită volumului mare de memorie pe care îl necesită, manipularea imaginilor cu organizare bitmap este dificilă, motiv pentru care au fost puse la punct **tehnici de compresie** ce reduc semnificativ volumul de memorie, fără a altera calitatea imaginii, din punctul de vedere al aplicaţiei. Pentru sistemele de operare din familia Windows dezvoltate de firma Microsoft sunt disponibile mai multe **standarde** de stocare a imaginilor digitale, care constituie totodată şi tipuri de fişiere, cum ar fi:

- **BMP** (Bit-Mapped Picture) – format de fişier de tip bitmap propriu pentru Microsoft, care stochează fără pierderi imagini având 32 de biţi pe pixel;
- **JPEG** (Joint Photographic Expert`s Group) – în prezent unul dintre cele mai larg utilizate formate de compresie pentru imagini color statice, disponibil pentru multe platforme software şi dispozitive hardware; poate suporta până la 24 biţi de culoare pe pixel; lucrează independent de sistemul de codare a culorii; deşi acţionează "cu pierderi", nivelul de degradare al imaginilor este controlabil şi, în multe cazuri, nedetectabil; compresia se bazează pe eliminarea informaţiilor care nu sunt uşor detectabile pentru ochiul uman, acţionând optim când imaginea prezintă regiuni mari de culori aproape constante şi când variaţiile de înaltă frecvenţă în regiunile de detaliu nu sunt importante pentru utilizator; eficienţa în stocarea şi transmisia imaginilor este foarte bună (fişierul comprimat poate avea 1/20 sau 1/30 din dimensiunea fişierului original);
- **TIFF** (Tagged Image File Format) – format de fişier foarte general şi foarte complex, care suportă imagini de tip bitmap, profunzimea maximă a culorii fiind de 24 de biţi; compresia se poate face cu sau fără pierderi;
- **PCX** – format de fişier de tip bitmap cu maxim 32 de biţi pe pixel, specific sistemului de operare MS-DOS;
- **GIF** (Graphics Interchange Format) – format de fişier frecvent utilizat pentru codarea imaginilor existente în World Wide Web sau în baze de date generale, acest lucru fiind determinat de raportul acceptabil precizie/dimensiune; suportă maxim 8 biţi de culoare pe pixel, motiv pentru care nu este utilizat în cazul imaginilor ce necesită o înaltă precizie a culorilor; oferă o dimensiune de fişier eficientă, în dauna calităţii imaginii.

2.1.4. CRITERII ÎN SELECTAREA CÂMPURILOR PENTRU IMAGINI DIGITIZATE DE ÎNALTĂ CALITATE

Imaginile analogice de microscopie optică ce urmează a fi supuse unei procesări de imagine asistate de calculator trebuie anume alese, spre a oferi posibilitatea analizei cantitative asupra unor componente structurale definitorii pentru respectivele tipuri de elemente morfologice. **Relevanţa** imaginilor digitizate rezultate prin achiziţia, cu o cameră video ataşată la un microscop, a câmpurilor microscopice selectate se decide în baza unei duble condiţii:

Scanned with OKEN Scanner

informaţia biomedicală, în sensul analizei calitative, şi **informaţia cuantificabilă** din punct de vedere matematic, care asigură premisele analizei cantitative.

Abordarea algoritmică, indispensabilă determinărilor automate ale unor parametri, impune anumite **proprietăţi** pentru imaginile digitizate (după cum se va detalia mai jos), în lipsa cărora rezultatele procesării morfologice cantitative, asistate de calculator, sunt, de obicei, afectate de erori majore. Astfel, am constatat că respectarea următoarelor criterii, privind organizarea informaţiei în câmpul microscopic, este crucială pentru asigurarea unei cât mai bune precizii în efectuarea determinărilor numerice pe imaginile digitizate, fără intervenţia factorului uman:

- prezenţa, în câmpul analizat, a unui număr rezonabil de structuri ce fac obiectul măsurătorilor, în sensul că: (*i*) prea puţine structuri în câmp pot conduce la concluzii cu valabilitate restrânsă, (*ii*) supraîncărcarea câmpului cu structuri revine (din punctul de vedere al imaginii digitizate) la un număr scăzut de pixeli corespunzând fiecărei structuri în parte şi, drept consecinţă, la diminuarea preciziei evaluărilor numerice;
- buna demarcare a structurilor diferite, prin existenţa de contururi cât mai precise şi/sau a diferenţelor de culoare;
- raportarea la un reper bidimensional local al imaginii din câmpul vizual, de aşa manieră încât imaginea digitizată să rezulte convenabil plasată în fereastra de lucru, în sensul posibilităţii de a efectua măsurători raportate la sistemul de coordonate;
- luminozitatea uniformă a câmpului, de aşa manieră încât imaginea digitizată să nu introducă reflexe şi umbre.

Aparent, neîndeplinirea parţială a acestor criterii poate fi compensată prin operaţii de preprocesare a imaginii digitale, sub programe ce oferă astfel de facilităţi (de exemplu programe de uz general – Adobe Photoshop (Adobe Systems, Inc., 1995), Microsoft PhotoEditor (Microsoft Corporation, 1996), sau programe specializate – Zeiss KS400 (Kontron Elektronik GmbH, 2002, Carl Zeiss Vision GmbH, 2002), OPTIMAS (Media Cybernetics Inc., 2002), IMAGIC-5 (Image Science Software GmbH, 2002), NIH Image (National Institutes of Health – Research Services Branch, 2002)). Atare intervenţii nu sunt totuşi recomandate drept practică curentă, deoarece efectul compensator scontat este însoţit, în majoritatea cazurilor, de alterarea informaţiei originale. Excepţie poate face neîndeplinirea celui de-al doilea criteriu, datorată, de exemplu, unei colorаţii mai slabe, sau cvasi-uniformităţii de culoare între citoplasma celulară şi ansamblul ţesuturilor, specifică metodei hematoxilin-eozină; în asemenea situaţii se poate obţine o certă îmbunătăţire prin ajustarea contrastului şi a luminozităţii, pe imaginea digitizată.

2.2. CONCEPTE ŞI INSTRUMENTE

După cum s-a constatat deja de la începutul capitolului curent, abordarea tratării imaginilor digitale a impus folosirea **limbajului propriu** domeniului procesării imaginilor. În acest sens, secţiunea de faţă este în totalitate dedicată introducerii unor noi **termeni specifici**, definind **concepte** şi **instrumente** care stau la baza dezvoltării aplicaţiilor practice în prelucrarea asistată de calculator a imaginilor. Prima subsecţiune are drept obiectiv schiţarea unei etapizări pentru activităţile care se succed până a intra în posesia rezultatelor măsurătorilor, activităţi referite în ansamblu sub numele generic de procesare (prelucrare) numerică (digitală) de imagine. În vederea detalierii ulterioare a acestor etape, am considerat oportun ca următoarele patru subsecţiuni să explice şi să ilustreze câteva noţiuni indispensabile pentru înţelegerea materialului din restul capitolului: histograma unei imagini, vecinătăţile unui pixel, elemente structurale sau măşti, regiuni într-o imagine digitală.

Scanned with OKEN Scanner

2.2.1. ETAPELE PROCESĂRII NUMERICE A IMAGINII

Odată ce dispunem de o imagine digitizată, aceasta poate face obiectul unei procesări (prelucrări) numerice care, în cazul domeniului morfologiei normale şi patologice, se finalizează cu o serie de informaţii de ordin cantitativ privitoare la entităţile prezente în imaginea respectivă. Succesiunea de operaţii care conduce la atare rezultate este, în general, suficient de complexă, definind un număr de **etape standard**, după cum urmează:

- **Preprocesarea imaginii (procesarea preliminară)** – are drept scop îmbunătăţirea calităţii imaginii digitale, crescând capacitatea de detecţie a unor detalii importante pentru ochiul uman şi/sau pentru procedeul automat implementat pe calculator.
- **Segmentarea** – permite a separa, în cadrul imaginii digitale, zone cu o anumită semnificaţie pentru observator, sau pentru derularea în ansamblu a unei aplicaţii implementate pe calculator.
- **Efectuarea de operaţii binare** – se referă la manipularea imaginii alb-negru ce poate rezulta în urma procesului de segmentare şi în care se află, sub formă de zone albe, corespondenţi ai unor structuri de interes din imaginea iniţială, monocromă sau color. Obiectivul manipulării constă în corectarea unor neajunsuri inerente etapei de segmentare.
- **Analiza cantitativă** – extrage, prin tehnici de măsurare adecvată, informaţia numerică ce caracterizează cantitativ entităţile din imagine, în conformitate cu anumite criterii.

Fiecare din etapele enunţate include, în principiu, mai multe tipuri de operaţii, care vor fi detaliate pe parcursul secţiunilor viitoare din acest capitol.

2.2.2. HISTOGRAMA ŞI UTILIZAREA EI ÎN CARACTERIZAREA IMAGINILOR DIGITALE

O caracterizare de tip global a unei imagini monocrome este realizată prin **histograma nivelurilor de gri**. Histograma unei imagini de scală gri de 8 biţi este un **grafic** care exprimă numărul de pixeli din imagine, corespunzător unei anumite **luminozităţi**, având pe ordonată numărul de pixeli şi pe abscisă cele 256 niveluri de gri (de la 0 la 255).

În baza convenţiei prezentate în subsecţiunea 2.1.3, valoarea 0 (cu semnificaţie de negru) se ataşează pixelilor cu grad minim de luminozitate, valoarea $255 = 2^8-1$ (cu semnificaţie de alb) se ataşează pixelilor cu grad maxim de luminozitate, iar pixelilor cu grad intermediar de luminozitate (adică gri propriu-zis, fără alb şi negru) li se atribuie valori din gama 1-254. Maximele din histogramă corespund valorilor de luminozitate mai frecvente care, adesea, identifică structuri particulare prezente într-o imagine microscopică, în timp ce minimele indică valori de luminozitate care sunt mai puţin frecvente. Existenţa în histogramă a unor regiuni cu ordonată 0, situate de obicei la ambele capete, indică faptul că nici un pixel nu are acele valori de gri, deci luminozitatea imaginii nu acoperă întregul domeniu disponibil, de 0-255.

În cele ce urmează vom ilustra, prin două exemple, modul în care trebuie interpretate informaţiile furnizate de o histogramă.

Exemplul 2.2.1

Se consideră imaginea din fig. 2.2.1, reprezentând un corpuscul renal. Imaginea este caracterizată printr-o luminozitate ridicată, iar histograma ei este dată în fig. 2.2.2. Histograma, împreună cu informaţiile numerice din tabelul alăturat, indică plasarea nivelurilor de gri între valorile 16 (L – minim, de la *eng.* low) şi

Scanned with OKEN Scanner

255 (H – maxim, de la *eng.* high), cu un maxim absolut la 255 (100% pe axa ordonatelor) şi un maxim relativ la 250 (aproximativ 60% pe axa ordonatelor). Aceste valori corespund numărului mare de pixeli albi sau gri foarte deschis din imagine. Pe intervalul [0-15] ordonata este nulă, reflectând faptul că în imagine nu există pixeli de culoare neagră (valoare 0), şi nici pixeli de culoare gri foarte închis (valori 1-15). Se observă că foarte puţini pixeli corespund nuanţelor de gri închis. Inclusiv pixelii cu valori de gri mediu sunt relativ puţini, pentru domeniul 50-200 histograma nedepăşind 20% din valoarea maximă.

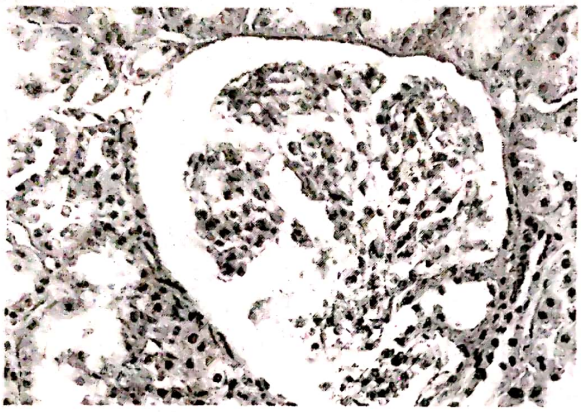
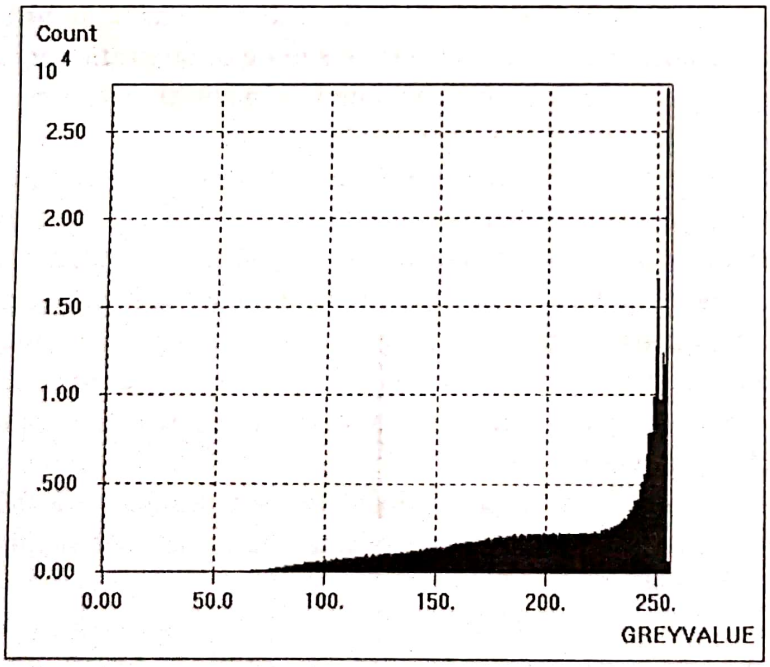


Fig. 2.2.1. Corpuscul renal – imagine caracterizată prin luminozitate ridicată.



Statistics	
file	2
parameter	GREYVALUE
counts	433637
lower bound	0
upper bound	255
underflow	0
in range	433637
overflow	0
classes	256
modul	1
min.count	0
max.count	27647
Statistics	
minimum	16
maximum	255
sum	8.74145e+007
medianvalue	216.113
mean	201.585
variance	2518.91
std.dev	50.1888
skewness	-0.846875
kurtosis	-0.257926

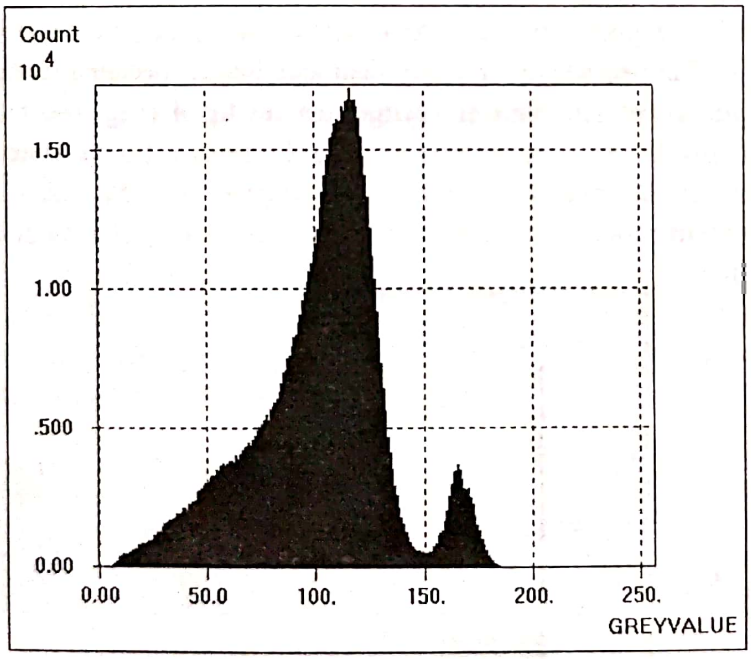
Fig. 2.2.2. Histograma de luminozitate corespunzătoare imaginii din fig. 2.2.1 şi tabelul cu informaţii statistice aferente.

Scanned with OKEN Scanner

Exemplul 2.2.2

Se consideră imaginea din fig. 2.2.3, reprezentând un ansamblu smalţ – dentină. Imaginea este caracterizată printr-o luminozitate scăzută, iar histograma ei este dată în fig. 2.2.4. Histograma indică plasarea nivelurilor de gri între 0 (L) şi 225 (H), cu un maxim absolut la 115 (100% pe axa ordonatelor) şi un maxim relativ la 165 (aproximativ 20% pe axa ordonatelor). Aceste valori corespund numărului mare de pixeli gri mediu şi gri închis din imagine. Totodată, din imagine lipsesc complet culoarea albă şi nuanţele de gri foarte deschis, neexistând nici un pixel în zona valorilor de gri 225-255, fapt reflectat de ordonata nulă pe acest interval.

Fig. 2.2.3. Ansamblu smalţ-dentină – imagine caracterizată prin luminozitate scăzută.



Statistics	
file	2
parameter	GREYVALUE
counts	908400
lower bound	0
upper bound	255
underflow	0
in range	908400
overflow	0
classes	256
modul	1
min.count	0
max.count	17344
Statistics	
minimum	0
maximum	225
sum	9.23948e+007
medianvalue	106.837
mean	101.712
variance	1007.56
std.dev.	31.7422
skewness	-0.317309
kurtosis	0.397669

Fig. 2.2.4. Histograma de luminozitate corespunzătoare imaginii din fig. 2.2.3 şi tabelul cu informaţii statistice aferente.

Într-o imagine digitală, fără a ţine cont de poziţia pixelilor în imagine, pot fi efectuate anumite operaţii care transformă luminozitatea pixelilor, conducând la modificări ale

histogramei, în sensul **deplasării**, **întinderii** (expandării) sau **egalizării** acesteia. Discutarea în detaliu a respectivelor modificări se va face pe parcursul subsecțiunilor 2.3.1 și 2.3.2, când vom aborda ajustarea luminozității și manipularea contrastului.

2.2.3. VECINĂTĂȚILE UNUI PIXEL

Pixelii dintr-o imagine digitală caracterizată *M*x*N* sunt organizați într-o **manieră tabelară**, corespunzător modului de dispunere a elementelor într-un tabel cu M coloane și N linii (pe fiecare din cele N linii orizontale ale imaginii sunt plasați M pixeli). Astfel, fiecărui pixel i se atribuie în mod unic o pereche de coordonate (i,j), cu $i = 0, \ldots, M-1$, și $j = 0, \ldots, N-1$. Convențional, se plasează pixelul de coordonate (0,0) în colțul din stânga sus a imaginii, ceea ce înseamnă că pixelii din celelalte trei colțuri vor avea coordonatele: $(0, M-1)$ – dreapta sus, $(0, N-1)$ – stânga jos, $(M-1, N-1)$ – dreapta jos.

Oricărui pixel din imagine i se poate atașa o mulțime de pixeli din imediata lui apropiere (cu care acesta se atinge), constituindu-se într-o entitate topologică denumită **vecinătate** (*eng.* connectivity). Operarea într-o manieră automatizată cu vecinătăți impune definirea vecinătăților într-un mod sistematizat. De exemplu, pentru orice pixel de coordonate (i,j), care nu aparține nici uneia din extremitățile imaginii, se poate considera mulțimea de pixeli cu coordonatele $(i-1,j-1)$, $(i-1,j)$, $(i-1,j+1)$, $(i,j+1)$, $(i+1,j+1)$, $(i+1,j)$, $(i+1,j-1)$, $(i,j-1)$ care constituie o vecinătate de 8 pixeli, cu reprezentarea grafică din fig. 2.2.5(a). O astfel de vecinătate este referită și cu denumirea de **vecinătate de tip 8** (*eng.* 8-connectivity), frecvent simbolizată V_8, considerându-se că toți cei 8 pixeli de coordonate enumerate anterior sunt conectați cu pixelul de coordonate (i,j); grafic vorbind, presupunând că fiecare pixel este reprezentat ca un pătrat, așa cum se sugerează prin fig. 2.2.5(a), pixelul (i,j) cu poziție centrală se atinge cu cei din vecinătate atât pe laturi, cât și pe colțuri. Din necesitatea de a face distincție între conectarea (atingerea) pe laturi și, respectiv, pe colțuri, a fost introdusă și **vecinătatea de tip 4** (*eng.* 4-connectivity), simbolizată V_4, care conține pixelii de coordonate $(i-1,j)$, $(i,j+1)$, $(i+1,j)$, $(i,j-1)$, conectarea cu pixelul de coordonate (i,j) având loc numai pe laturi (fig. 2.2.5(b)). Totodată, menționăm că, dacă vrem să ne referim la pixelii conectați pe colțuri, ei sunt caracterizați global ca diferență de mulțimi $V_8 \setminus V_4$ (fig. 2.2.5(c)).

$(i-1,j-1)$	$(i-1,j)$	$(i-1,j+1)$
$(i,j-1)$	(i,j)	$(i,j+1)$
$(i+1,j-1)$	$(i+1,j)$	$(i+1,j+1)$

Fig. 2.2.5(a). Reprezentarea grafică a vecinătății de tip 8.

	$(i-1,j)$	
$(i,j-1)$	(i,j)	$(i,j+1)$
	$(i+1,j)$	

Fig. 2.2.5(b). Reprezentarea grafică a vecinătății de tip 4.

Scanned with OKEN Scanner

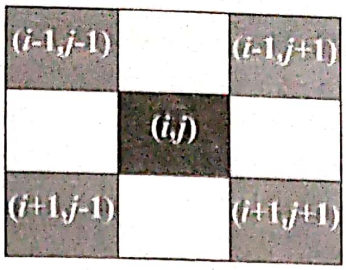


Fig. 2.2.5(c). Reprezentarea grafică a diferenței de mulțimi $V_8 \setminus V_4$.

2.2.4. ELEMENTE STRUCTURALE SAU MĂŞTI

Algoritmii folosiți în operațiile specifice procesării de imagini utilizează, pentru a explora imaginile digitale, mulțimi de poziții de pixeli învecinați denumite **elemente structurale**, **nuclee** (*eng.* kernels) sau **măşti** (*eng.* masks). Elementele structurale pot conține un număr arbitrar de pixeli, plasați pe poziții arbitrare, ceea ce înseamnă mărimi şi/sau forme diferite. De exemplu, măştile pot avea forme convenționale (de dreptunghi, pătrat, disc, inel) sau forme neconvenționale (în L, în M).

Pixelul de interes, denumit şi **origine**, poate fi oricare dintre pixelii elementului structural, dar cel mai adesea este pixelul central (valabil numai în cazul elementelor structurale simetrice). În cazul general, fiecărui pixel dintr-o mască îi pot fi atribuite ponderi cu valori diferite de la pixel la pixel. Pe parcursul acestui material nu vom face însă apel la astfel de abordări generale, ci vom considera măşti cu toți pixelii având pondere 1.

Fiecare produs software are propria lui gamă de elemente structurale aplicabile. Cel mai frecvent sunt însă folosite **modelele simetrice**, de tip pătrat (3x3, 5x5 etc.). Spre exemplificare, fig. 2.2.6 prezintă elemente structurale formate din 3x3 pixeli, care operează pe pixelul central (simbolizat cu negru), folosind informația provenită de la anumiți pixeli vecini, cu dispoziție simetrică (simbolizați cu gri). În cazul (a), informația de la toți cei 8 pixeli adiacenți este aplicată în determinarea valorii pixelului central. În cazul (b) sunt utilizați numai 4 pixeli adiacenți, şi anume cei situați pe verticală şi pe orizontală. În cazurile (c) şi (d) sunt utilizați numai 2 pixeli adiacenți, şi anume cei situați pe orizontală (c), respectiv pe verticală (d). În cazurile (e) şi (f) sunt, de asemenea, utilizați numai 2 pixeli vecini, şi anume cei situați pe diagonala corespunzătoare primei bisectoare (e) şi, respectiv, pe diagonala corespunzătoare celei de a doua bisectoare (f).

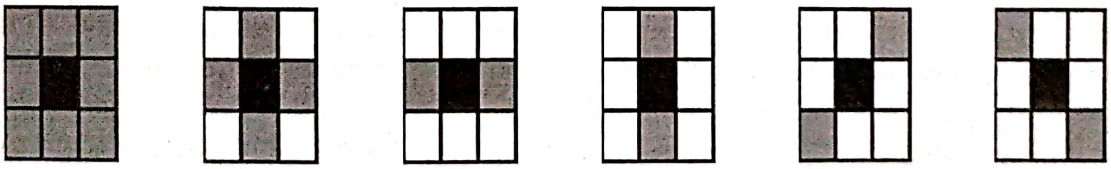


Fig. 2.2.6. Elemente structurale (măşti) formate din 3x3 pixeli, cu pixelul de interes central. Modalitate de operare diferențiată, în raport cu informația provenită de la anumiți pixeli vecini.

Imaginea inițială este examinată pixel cu pixel, aplicându-se un anumit tip de element structural, astfel încât originea măştii să corespundă pixelului de interes. Prin alegerea unui

Scanned with OKEN Scanner

anumit element structural se stabileşte care dintre pixelii din imediata apropiere a pixelului de interes (din imaginea iniţială) sunt utilizaţi pentru a determina valoarea pixelului cu aceleaşi coordonate din imaginea rezultată. Altfel spus, se realizează o nouă valoare (în sensul de culoare, nuanţă de gri, valoare binară – alb sau negru) pentru pixelul de interes, bazată pe valoarea pixelilor înconjurători.

2.2.5. REGIUNI ÎNTR-O IMAGINE DIGITALĂ

Regiunea (*eng.* region) (în sensul strict al limbajului tehnic de specialitate) reprezintă o mulţime de pixeli conectaţi între ei prin raportare la un acelaşi tip de vecinătate. Astfel, considerând cele două tipuri de vecinătăţi prezentate în subsecţiunea anterioară şi folosind simbolizarea grafică ce asociază fiecărui pixel un pătrat, putem întâlni cele două situaţii din fig. 2.2.7. În fig. 2.2.7(a), dacă pentru pixelul P1 se consideră o vecinătate de tip V_8, atunci pixelul P2 aparţine acestei vecinătăţi şi reciproc, ceea ce înseamnă că pixelii P1 şi P2 aparţin unei aceleiaşi regiuni, prin raportare la vecinătatea de tip V_8. Tot în fig. 2.2.7(a), dacă pentru pixelul P1 se consideră o vecinătate de tip V_4, atunci pixelul P2 nu aparţine acestei vecinătăţi şi reciproc, ceea ce înseamnă că pixelii P1 şi P2 nu aparţin unei aceleiaşi regiuni (P1 aparţine unei regiuni, P2 unei alte regiuni), prin raportare la vecinătatea de tip V_4. Spre deosebire de fig. 2.2.7(a), în fig. 2.2.7(b) pixelii P1 şi P2 sunt conectaţi între ei prin raportare la ambele tipuri vecinătăţi (V_8 şi V_4), ceea ce înseamnă că pixelii P1 şi P2 aparţin unei aceleiaşi regiuni, indiferent de tipul vecinătăţii considerate.

Fig. 2.2.7. Încadrarea pixelilor în regiuni: (a) o singură regiune în raport cu V_8 sau două regiuni în raport cu V_4; (b) o singură regiune în raport cu V_8 sau cu V_4.

Identificarea **automată** a regiunilor se realizează în etapa de segmentare (după cum s-a precizat în subsecţiunea 2.2.1 şi se va detalia în secţiunea 2.4). O altă manieră de lucru care conduce la definirea regiunilor este cea **interactivă**, constând în desenarea de către utilizator a contururilor acestora. În cazul particular al imaginilor binare (alb-negru) există două tipuri de regiuni: **găuri** (alcătuite exclusiv din pixeli negri, cu valoare de gri 0) şi **obiecte** (alcătuite exclusiv din pixeli albi, cu valoare de gri 255).

Importanţa conceptului de regiune rezultă din faptul că procesul de măsurare, corespunzător etapei de analiză cantitativă, se bazează pe regiuni. Diferite platforme software au definite **seturi de proprietăţi** specifice conceptului de regiune, care pot fi aplicate (ca atare sau adaptate) în efectuarea măsurătorilor. În cazul general, proprietăţile unei regiuni pot fi clasificate în două mari categorii, şi anume **geometrice** (care nu ţin cont de valoarea de culoare sau nuanţă de gri a pixelilor) şi **densitometrice/texturale** (care ţin cont de valoarea de culoare sau nuanţă de gri a pixelilor). Având în vedere specificul măsurătorilor practicate în mod curent în histologie şi histopatologie, proprietăţile care fac obiectul analizei cantitative sunt, în principal, de natură geometrică. Într-o imagine, de obicei sunt identificate mai multe

Scanned with OKEN Scanner

regiuni. Mediile software specializate ţin automat evidenţa acestora şi efectuează măsurătorile solicitate de utilizator asupra fiecărei regiuni în parte.

În strânsă legătură cu noţiunea de regiune apare obligatorie definirea conceptului de **câmp**. Câmpul reprezintă întreaga imagine supusă procesării sau, în anumite situaţii, o porţiune din aceasta, delimitată de un contur anterior definit; câmpul conţine regiuni. Realizate la nivel de câmp, măsurătorile (focalizate pe trăsături specifice câmpului) caracterizează global întreaga imagine.

2.3. PROCESAREA PRELIMINARĂ – TEHNICI DE ÎMBUNĂTĂŢIRE A IMAGINILOR DIGITALE

Un avantaj semnificativ al imaginilor digitale faţă de cele analogice este faptul că primele pot fi uşor manipulate în scopul creşterii vizibilităţii unor componente sau a unor aspecte dintr-o structură de interes.

Îmbunătăţirea (*eng.* enhance) unei imagini se bazează pe o serie de tehnici ce urmăresc a creşte calitatea aspectului vizual, sau a converti imaginea într-o formă mai potrivită pentru interpretarea umană sau computerizată. În prezent, *nu există încă o teorie generală a îmbunătăţirii imaginii, deoarece nu există un standard general pentru calitatea unei imagini*. Astfel, în scopul de a accentua unele trăsături ale imaginii, pentru afişarea informaţiei vizuale sau pentru extragerea unor caracteristici, în ultimele decenii au fost dezvoltate diferite tipuri de **tehnici**, destinate imaginilor color sau cu niveluri de gri: ajustarea luminozităţii, manipularea contrastului, îndepărtarea zgomotelor şi umbrelor, detecţia marginilor, identificarea contururilor şi separarea regiunilor. Toate aceste tehnici urmează a fi discutate în continuare, prezentarea teoretică fiind însoţită de exemple adecvate.

2.3.1. AJUSTAREA LUMINOZITĂŢII

Ajustarea luminozităţii unei imagini înseamnă **modificarea histogramei** asociate acesteia, prin **translare la dreapta** (pentru creşterea luminozităţii) şi, respectiv, **la stânga** (pentru reducerea luminozităţii), fără a afecta alura graficului reprezentat de histograma iniţială. Aceasta revine, pentru o imagine monocromă cu 256 niveluri de gri, la modificarea luminozităţii fiecărui pixel din imagine (notată generic prin x) conform formulei:

$$y(x)=\begin{cases} x+C, & x\in[0,255-C); \\ 255, & x\in[255-C,255]. \end{cases} \quad 1\le C\le 254 \tag{2.3.1a}$$

în cazul deplasării spre dreapta a histogramei, şi, respectiv, a formulei:

$$y(x)=\begin{cases} 0, & x\in[0,C); \\ x-C, & x\in[C,255]. \end{cases} \quad 1\le C\le 254 \tag{2.3.1b}$$

în cazul deplasării spre stânga a histogramei.

Merită de subliniat faptul că, în urma deplasării la dreapta, imaginea modificată nu va conţine nuanţele de gri din intervalul [0,C-1], iar în urma deplasării la stânga, imaginea modificată nu va conţine nuanţele de gri din intervalul [256-C,255]. Astfel, pentru o valoare rezonabilă a constantei C de ajustare a luminozităţii, în urma deplasării la dreapta, imaginea modificată nu va conţine nuanţele de gri foarte închis şi negru, iar în urma deplasării la stânga, imaginea modificată nu va conţine nuanţele de gri foarte deschis şi alb.

Scanned with OKEN Scanner

Exemplul 2.3.1

Se consideră imaginea din fig. 2.3.1, reprezentând teaca Hertwig în dezvoltarea mugurelui dentar. Imaginea este caracterizată prin niveluri de gri între 55 (L) si 145 (H), însemnând un domeniu de 90 de niveluri de gri, conform histogramei din fig. 2.3.2. Imaginea din fig. 2.3.3 este rezultatul creşterii luminozităţii prin adăugarea valorii $C = 70$ la nivelul de gri al fiecărui pixel din imaginea iniţială, conform formulei (2.3.1a), de deplasare a histogramei la dreapta. În consecinţă, histograma din fig. 2.3.4 este deplasată spre zona de alb cu 70 de unităţi, ocupând un domeniu tot de 90 de niveluri de gri, situat între nivelurile de gri 125 (L) – 215 (H) şi păstrând aceeaşi formă ca în cazul imaginii iniţiale.

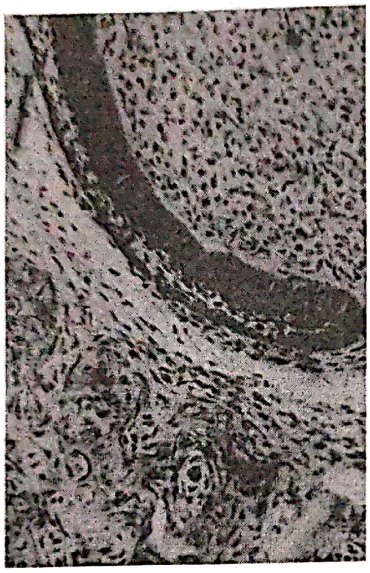
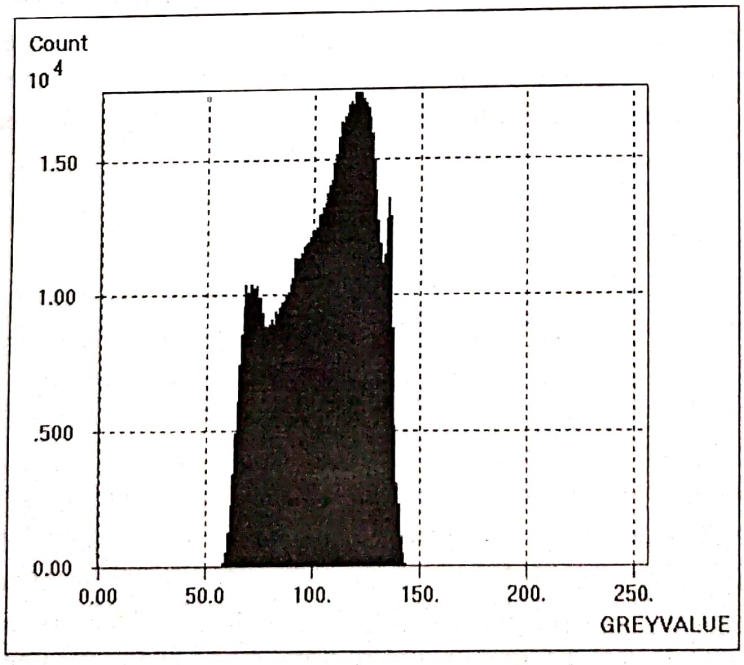


Fig. 2.3.1. Mugure dentar – teaca Hertwig – imagine iniţială pentru procedeul de ajustare a luminozităţii.



Statistics	
file	2
parameter	GREYVALUE
counts	955200
lower bound	0
upper bound	255
underflow	0
in range	955200
overflow	0
classes	256
modul	1
min.count	0
max.count	17592
Statistics	
minimum	56
maximum	145
sum	9.94231e+007
medianvalue	107.502
mean	104.086
variance	443.581
std.dev.	21.0614
skewness	-0.290907
kurtosis	-1.01612

Fig. 2.3.2. Histograma de luminozitate corespunzătoare imaginii din fig. 2.3.1 (acoperind un domeniu de 90 de niveluri de gri) şi tabelul cu informaţii statistice aferente.

Scanned with OKEN Scanner

Fig. 2.3.3. Imagine rezultată ca urmare a creşterii luminozităţii prin translarea la dreapta a histogramei iniţiale.

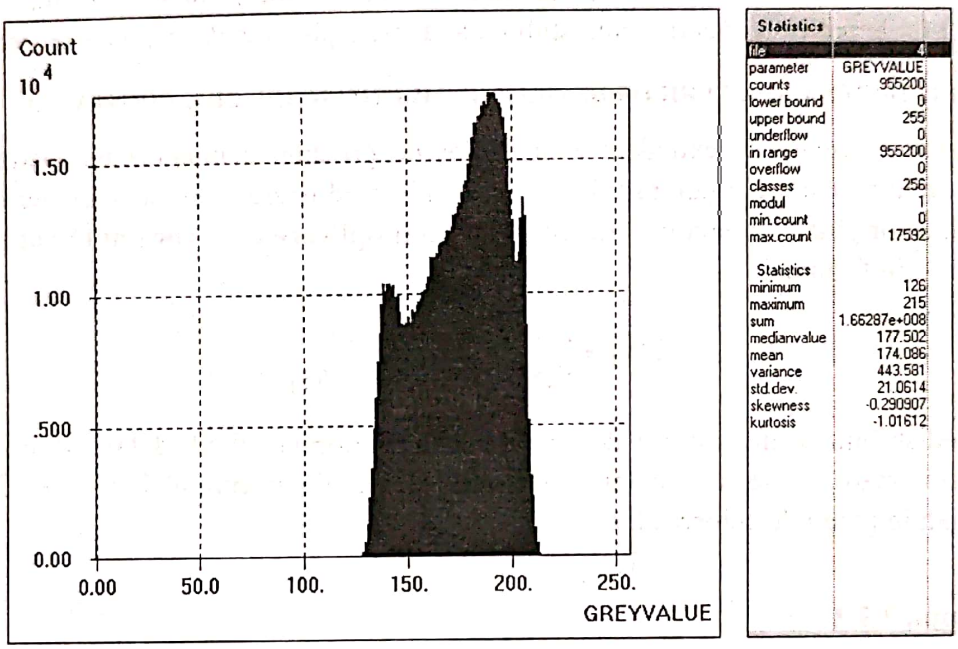


Fig. 2.3.4. Histograma de luminozitate corespunzătoare imaginii din fig. 2.3.3 şi tabelul cu informaţii statistice aferente.

2.3.2. MANIPULAREA CONTRASTULUI

Ajustarea luminozităţii nu este un procedeu performant, în sensul îmbunătăţirii calităţii imaginilor, deoarece histograma îşi păstrează alura şi, implicit, va păstra o zonă îngustă în domeniul nuanţelor de gri folosite. Creşterea performanţelor ar necesita ca histograma rezultată în urma modificărilor să ocupe o gamă mult mai largă de nuanţe de gri, ceea ce nu se poate realiza printr-o simplă translare a graficului. Un procedeu ce răspunde la o atare necesitate este **manipularea contrastului**, care poate fi realizată în mai multe maniere. Acest procedeu este în strânsă legătură cu modul în care sunt percepute imaginile de către ochiul uman, corespunzând, totodată, şi cerinţelor impuse de metodele automate de prelucrare a imaginilor.

Scanned with OKEN Scanner

Calitatea percepției ochiului uman poate fi discutată, pornind de la impresia pe care o creează imaginile realizate corect prin fotografiere sau preluate cu camere video, la microscoapele de laborator. Deşi, conform literaturii, astfel de imagini utilizează numai 10-30 niveluri de gri, ele apar clare pentru observatorul uman. Impresia de claritate se justifică prin aceea că ochiul uman poate distinge doar câteva zeci de niveluri de gri, ceea ce înseamnă că el nu va percepe diferența dintre două nuanțe de gri învecinate dintr-un domeniu cu 256 de valori.

Îmbunătățirea contrastului urmăreşte să ajute percepția de către observatorul uman a unei imagini digitale monocrome, prin **redistribuirea nivelurilor de gri** a pixelilor pe o gamă mai largă, fără a lua în discuție informațiile legate de poziția pixelilor, în raport cu cele două coordonate. Problema este similară şi în cazul unei imagini color, pentru care s-a precizat că tehnica digitală poate pune la dispoziție până la milioane de culori.

Termenul de "manipulare a contrastului" include modificări atât în sensul **creşterii**, cât şi al **descreşterii**. Precizăm că, în cele ce urmează, ne vom referi explicit doar la metode care asigură **creşterea contrastului**, deoarece ne interesează îmbunătățirea calității imaginii, din punctul de vedere al observatorului uman. Evident, într-o manieră similară se poate aborda şi problematica reducerii contrastului, dar aceasta prezintă doar interes teoretic.

- **REDISTRIBUIREA NIVELURILOR DE GRI PRIN MULTIPLICARE CU O CONSTANTĂ**

O modalitate imediată de extindere a nuanțelor de gri dintr-o imagine monocromă, deci, implicit, de îmbunătățire a contrastului, o constituie modificarea luminozității fiecărui pixel din imaginea originală, notată generic cu x, prin **multiplicare** cu o **constantă supraunitară** $K > 1$, conform formulei:

$$y(x) = \begin{cases} Kx, & x \in [0, 255/K); \\ 255, & x \in [255/K, 255]. \end{cases} \qquad (2.3.2)$$

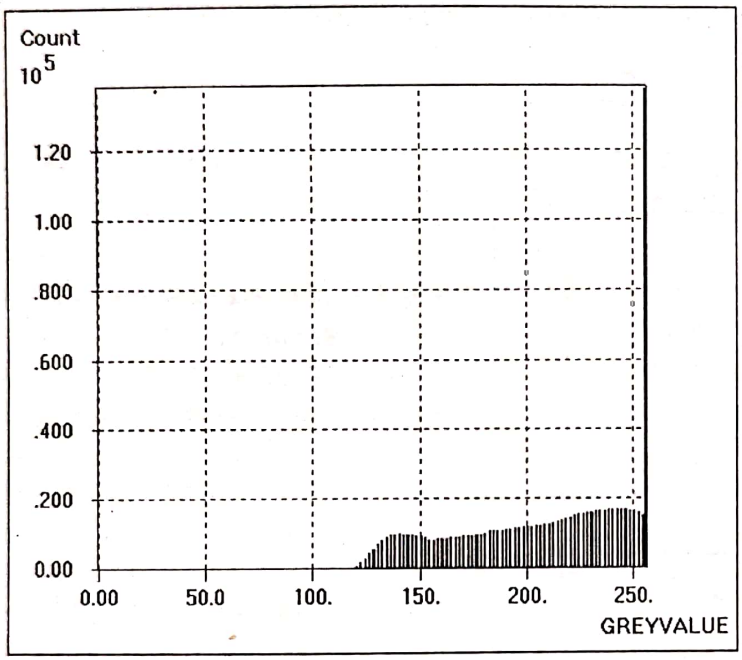
Redistribuirea nivelurilor de gri nu este uniformă, fiind defavorizată zona de luminozitate mare, în sensul că toți pixelii cu nuanțe din intervalul [255 / K, 255] sunt transformați în pixeli de culoare albă.

Exemplul 2.3.2

Se consideră imaginea din fig. 2.3.1, cu histograma din fig. 2.3.2, şi se urmăreşte îmbunătățirea contrastului, prin multiplicarea cu K = 2 a nivelului de luminozitate a fiecărui pixel (conform formulei (2.3.2)). Rezultă imaginea din fig. 2.3.5, a cărei histogramă este prezentată în fig. 2.3.6. Domeniul nivelurilor de gri în fig. 2.3.6 porneşte de la 55 x 2 = 110 şi ar trebui să ocupe 90 x 2 = 180 de niveluri de gri. Aceasta ar însemna ca domeniul să ajungă la un nivel maxim de 290, ceea ce este imposibil. Din acest motiv, în histograma din fig. 2.3.6 se constată dilatarea domeniului 55 – 127 din histograma inițială şi asocierea valorii maxime 255 (alb) tuturor pixelilor cu luminozitate mai mare de 127 în histograma inițială. O comparație între fig. 2.3.5 şi fig. 2.3.1 evidențiază îmbunătățirea contrastului, concomitent cu prezența masivă a culorii albe.

Scanned with OKEN Scanner

Fig. 2.3.5. Imagine rezultată ca urmare a îmbunătăţirii contrastului, datorită redistribuirii nivelurilor de gri prin multiplicare cu o constantă.



Statistics	
file	5
parameter	GREYVALUE
counts	955200
lower bound	0
upper bound	255
underflow	0
in range	955200
overflow	0
classes	256
modul	1
min.count	0
max.count	138025
Statistics	
minimum	112
maximum	255
sum	1.97347e+008
medianvalue	214.502
mean	206.603
variance	1601.01
std.dev.	40.0127
skewness	-0.44172
kurtosis	-1.04983

Fig. 2.3.6. Histograma de luminozitate corespunzătoare imaginii din fig. 2.3.5 şi tabelul cu informaţii statistice aferente.

- **REDISTRIBUIREA NIVELURILOR DE GRI PRIN ÎNTINDEREA (EXPANDAREA) HISTOGRAMEI**

Neajunsurile semnalate în procedeul de accentuare a contrastului expus anterior pot fi eliminate dacă **redistribuirea nuanţelor de gri** se realizează într-un **mod uniform**. Astfel, dacă în imaginea originală nivelurile de gri ocupă intervalul $[a, b]$ şi se doreşte ca, prin modificarea contrastului, nivelurile de gri să fie redistribuite în intervalul mai larg $[c, d]$, luminozitatea pixelilor (notată generic cu y) din noua imagine va rezulta din luminozitatea pixelilor (notată generic cu x) din imaginea iniţială, prin formula:

$$y(x) = \frac{d-c}{b-a}(x-a)+c\,,\ a \le x \le b\,,\ c \le y \le d\,. \qquad (2.3.3)$$

Scanned with OKEN Scanner

În cazul când intervalul $[a, b]$ este inclus în intervalul $[c, d]$, operația conduce doar la o **accentuare** a contrastului. O situație frecventă o constituie accentuarea contrastului în imaginile monocrome, pentru care intervalul $[a, b]$ se expandează la intervalul $[c, d] = [0, 255]$, formula de mai sus particularizându-se astfel:

$$y(x) = 255\ (x-a)/(b-a). \tag{2.3.4}$$

Exemplul 2.3.3

Se urmărește îmbunătățirea contrastului, pornind tot de la imaginea inițială din fig. 2.3.1 (ca și în exemplul precedent), domeniul tuturor nuanțelor de gri din imagine [55, 145] fiind transformat în domeniul [0, 255]. Procedeul s-a bazat pe aplicarea formulei (2.3.4), cu valorile concrete $y(x) = 255\ (x-55)/(145-55)$. Această funcție are graficul din fig. 2.3.7 și definește o transformare liniară a domeniului [55, 145] la domeniul [0, 255]. Imaginea din fig. 2.3.8, cu histograma din fig. 2.3.9, constituie rezultatul obținut prin aplicarea transformării. Astfel, noua histogramă ocupă întreg domeniul valorilor de gri [0, 255], dar păstrează alura celei originale, deformată aferent prin expandare.

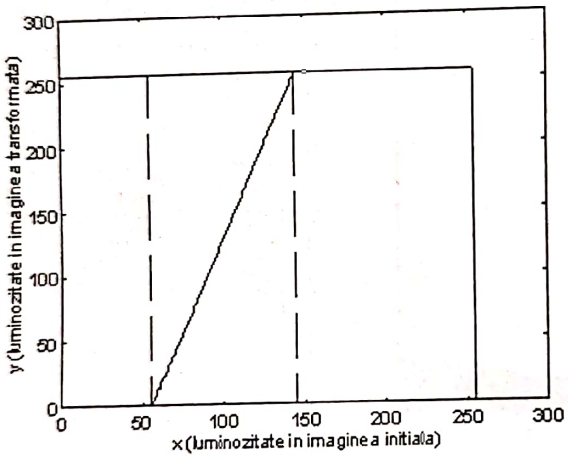
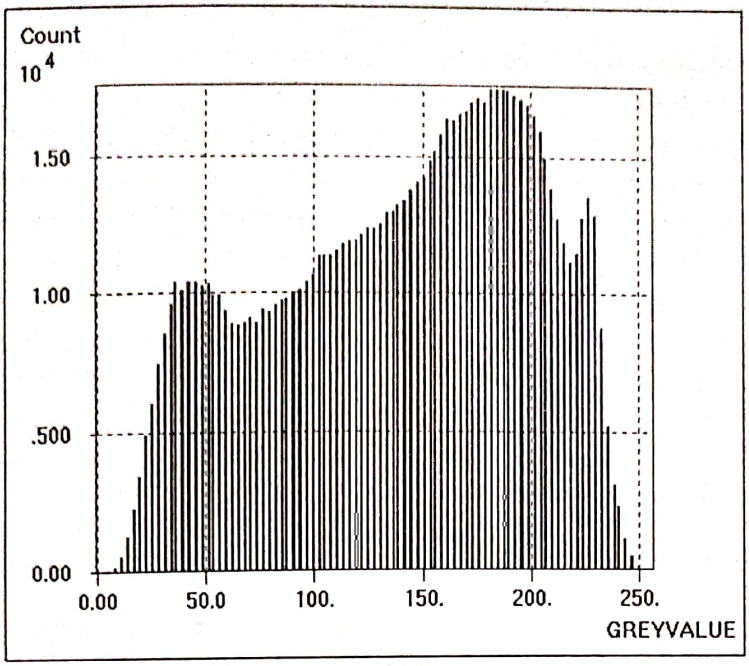


Fig. 2.3.7. Reprezentarea grafică a transformării liniare $y(x) = 255\ (x-55)/(145-55)$ aplicată histogramei de luminozitate.

Fig. 2.3.8. Imagine rezultată ca urmare a îmbunătățirii contrastului, datorită redistribuirii nivelurilor de gri prin expandarea histogramei.

Scanned with OKEN Scanner



Statistics	
file	G
parameter	GREYVALUE
counts	955200
lower bound	0
upper bound	255
underflow	0
in range	955200
overflow	0
classes	256
modul	1
min count	0
max count	17592
Statistics	
minimum	2
maximum	255
sum	1.32446e+008
medianvalue	147.502
mean	138.658
variance	3560.66
std.dev.	59.6713
skewness	-0.291099
kurtosis	-1.01615

Fig. 2.3.9. Histograma de luminozitate corespunzătoare imaginii din fig. 2.3.8 şi tabelul cu informaţii statistice aferente.

În cazul în care intervalul $[c, d]$ este ales astfel încât $a<c$ şi $b<d$, operaţia definită anterior are drept efect atât accentuarea contrastului (deoarece s-a precizat anterior că intervalul $[c, d]$ este mai larg decât intervalul $[a, b]$), cât şi creşterea luminozităţii (deoarece intervalul $[c, d]$ este situat la dreapta valorii a). Se observă cu uşurinţă că particularizarea acestui caz cu intervalele $[a, b]$ şi $[c, d]$ de lungime egală înseamnă doar creşterea luminozităţii (fără intervenţii asupra contrastului), formula de calcul reducându-se la

$$y(x)=x-a+c=x+(c-a),\ 1\le c-a\le 254, \qquad (2.3.5)$$

care coincide cu modificarea luminozităţii prin deplasarea histogramei la dreapta cu $C=c-a$ (vezi formula (2.3.1a)).

Exemplul 2.3.4

Considerăm situaţia de creştere a luminozităţii discutată în exemplul 2.3.1. Această situaţie poate fi interpretată şi prin prisma formulei (2.3.5), în care intervalul $[a, b] = [55, 145]$ este transformat în intervalul de lungime egală $[c, d] = [125, 215]$, ceea ce înseamnă o deplasare a histogramei la dreapta cu valoarea $C=c-a=125-55=70$.

După cum s-a constatat, procedeul de accentuare a contrastului prezentat în exemplul 2.3.3 expandează domeniul de valori de gri din intervalul $[a, b]$ printr-o **transformare liniară**. Există situaţii când această expandare se efectuează prin **transformări neliniare**, care au drept efect dilatări neuniforme ale subintervalelor din $[a, b]$, însemnând o tratare preferenţială a anumitor zone de gri.

Scanned with OKEN Scanner

Exemplul 2.3.5

Considerăm situația de accentuare a contrastului discutată în exemplul 2.3.2, care s-a bazat pe redistribuirea nivelurilor de gri, prin multiplicare cu o constantă. Această situație poate fi interpretată ca o transformare neliniară a domeniului de valori de gri $[a, b] = [55, 145]$, la întreg domeniul $[0, 255]$. Subintervalul $[55, 127]$ este expandat printr-o transformare liniară de pantă $K = 2$ în subintervalul $[110, 254]$, iar nivelurilor de gri din subintervalul $[128, 145]$ li se aplică o transformare liniară de pantă 0, prin care tuturor li se atașează valoarea constantă 255. În aceste condiții, transformarea de ansamblu, definită prin cele două transformări liniare de pantă diferită, este evident neliniară la nivelul întregului domeniu, după cum reiese și din fig. 2.3.10.

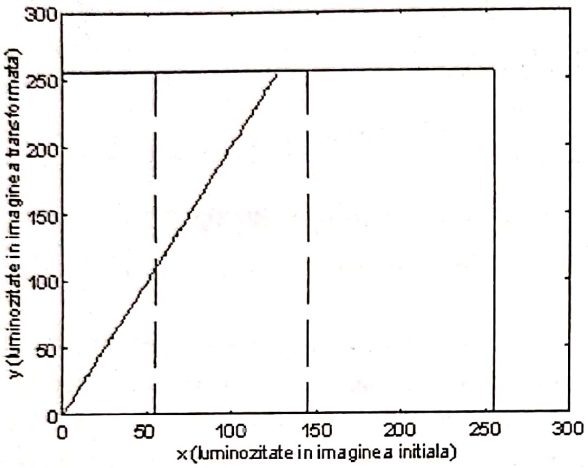


Fig. 2.3.10. Reprezentarea grafică a transformării neliniare

$$y(x) = \begin{cases} 2x, & x \in [55,127] \\ 255, & x \in [128,145] \end{cases}$$

aplicată histogramei de luminozitate.

- **REDISTRIBUIREA NIVELURILOR DE GRI PRIN EGALIZAREA HISTOGRAMEI**

Un procedeu care modifică semnificativ histograma și, implicit, calitatea imaginii ce rezultă din transformare are în vedere reducerea diferențelor dintre valorile maxime și respectiv, minime, în sensul diminuării "denivelărilor" în aspectul grafic al înfășurătoarei histogramei modificate. Din acest motiv, procedeul este cunoscut sub numele de **egalizarea histogramei**. El este recomandabil în cazul imaginilor în care este utilizat întregul domeniu de niveluri de gri, dar multe dintre aceste niveluri au asociate un număr redus de pixeli.

În histograma imaginii inițiale, vom nota prin $i = 0,1,\cdots,n-1$, valorile de gri utilizate și prin p_i numărul de pixeli corespunzător fiecăreia din aceste valori. De asemenea, vom nota prin N numărul total de pixeli din imagine, adică $N = \sum_{i=0}^{n-1} p_i$.

Ideal vorbind, în urma aplicării acestui procedeu, noua histogramă ar trebui să conțină un același număr de pixeli $q = N/n$, pe fiecare nivel de gri din gama $0,1,\cdots,n-1$. Acest fapt nu este posibil, deoarece în imaginea inițială cel puțin unul din niveluri are numărul de pixeli asociat mai mare decât q. De aceea, în practică, se urmărește ca redistribuirea să se realizeze în sensul plasării pe fiecare nivel de gri a unui număr de pixeli cât mai apropiat de valoarea q. Evident, această remarcă privește nivelurile de gri cu puțini pixeli din imaginea inițială (adică $p_i < q$). În cazul contrar, când un nivel de gri din histograma inițială prezintă foarte

mulți pixeli (adică $p_i > q$), acest număr de pixeli trebuie păstrat ca atare, pentru că nu există nici un criteriu în baza căruia să se realizeze o partajare a lui p_i pe mai multe niveluri de gri.

Modul de distribuire a nivelurilor de gri pe care îl asigură imaginea transformată permite vizualizarea unor detalii în regiuni care apăreau cu aspect aproape uniform în imaginea originală. De aceea, egalizarea histogramei este considerată o metodă puternică de îmbunătățire a imaginilor cu contrast scăzut. Egalizarea histogramei poate fi interpretată drept o transformare neliniară a histogramei de luminozitate, pentru a cărei prezentare preferăm o **formulare algoritmică iterativă**, pe etape, în locul unei exprimări analitice.

Etapa 1

Investigația histogramei imaginii inițiale începe de la nivelul de gri 0 către valori mai mari. Plasarea pixelilor în histograma transformată începe de la nivelul de gri 0 către valori mai mari.

Se determină cel mai mic întreg pozitiv m_1 cu proprietatea că dubla inecuație

$$m_1 q - \frac{1}{2}q \le \sum_{i=0}^{x} p_i \le m_1 q + \frac{1}{2}q \tag{2.3.6}$$

are soluții $x \in \{0,1,\ldots,n-1\}$. Să notăm cu C_1 mulțimea tuturor acestor soluții și să desemnăm prin c_1 elementul din C_1 pentru care suma $\sum_{i=0}^{x} p_i$ are valoarea cea mai apropiată de $m_1 q$, adică:

$$\left| \sum_{i=0}^{c_1} p_i - m_1 q \right| = \min_{x \in C_1} \left| \sum_{i=0}^{x} p_i - m_1 q \right|. \tag{2.3.7}$$

Introducând notația $q_1 = \sum_{i=0}^{c_1} p_i$, avem semnificația că pe primele m_1 niveluri de gri din histograma transformată (adică pe nivelurile $0,1,\ldots,z_1 = m_1 - 1$) trebuie plasați cei q_1 pixeli care corespund nivelurilor de gri $0,1,\ldots,c_1$ din histograma inițială. Plasarea acestor q_1 pixeli se va realiza pe un unic nivel de gri al histogramei transformate, și anume $y_1 = \left[\frac{z_1}{2}\right]$, unde [*] notează partea întreagă a numărului *.

Etapa 2

Investigația histogramei imaginii inițiale începe de la nivelul de gri $c_1 + 1$ către valori mai mari. Plasarea pixelilor în histograma transformată începe de la nivelul de gri $z_1 + 1$ către valori mai mari.

Cu m_1 și c_1 cunoscuți din *Etapa 1*, se determină cel mai mic întreg pozitiv m_2 cu proprietatea că dubla inecuație

$$m_1 q + m_2 q - \frac{1}{2}q \le \sum_{i=0}^{x} p_i \le m_1 q + m_2 q + \frac{1}{2}q \tag{2.3.8}$$

are soluții $x \in \{c_1 + 1,\ldots,n-1\}$. Să notăm cu C_2 mulțimea tuturor acestor soluții și să desemnăm prin c_2 elementul din C_2 pentru care suma $\sum_{i=0}^{x} p_i$ are valoarea cea mai apropiată de $(m_1 + m_2)q$, adică:

Scanned with OKEN Scanner

$$\left|\sum_{i=0}^{c_2} p_i - (m_1+m_2)q\right| = \min_{x\in C_2}\left|\sum_{i=0}^{x} p_i - (m_1+m_2)q\right|. \qquad (2.3.9)$$

Introducând notația $q_2 = \sum_{i=c_1+1}^{c_2} p_i$, avem semnificația că pe următoarele m_2 niveluri de gri din histograma transformată (adică pe nivelurile $z_1+1,\ldots,z_2 = z_1+m_2$) trebuie plasați cei q_2 pixeli care corespund nivelurilor de gri $c_1+1,\ldots,c_2$ din histograma inițială. Plasarea acestor q_2 pixeli se va realiza pe un unic nivel de gri al histogramei transformate, și anume $y_2 = \left[\frac{z_1+z_2}{2}\right]$.

Etapa k (de formă generală)

Investigația histogramei imaginii inițiale începe de la nivelul de gri $c_{k-1}+1$ către valori mai mari. Plasarea pixelilor în histograma transformată începe de la nivelul de gri $z_{k-1}+1$ către valori mai mari.

Cu $m_1,\ldots,m_{k-1}$ și $c_1,\ldots,c_{k-1}$ cunoscuți din etapele precedente, se determină cel mai mic întreg pozitiv m_k cu proprietatea că dubla inecuație

$$(m_1+\ldots+m_{k-1})q + m_k q - \frac{1}{2}q \le \sum_{i=0}^{x} p_i \le (m_1+\ldots+m_{k-1})q + m_k q + \frac{1}{2}q \qquad (2.3.10)$$

are soluții $x\in\{c_{k-1}+1,\ldots,n-1\}$. Să notăm cu C_k mulțimea tuturor acestor soluții și să desemnăm prin c_k elementul din C_k pentru care suma $\sum_{i=0}^{x} p_i$ are valoarea cea mai apropiată de $(m_1+\ldots+m_k)q$, adică:

$$\left|\sum_{i=0}^{c_k} p_i - (m_1+\ldots+m_k)q\right| = \min_{x\in C_k}\left|\sum_{i=0}^{x} p_i - (m_1+\ldots+m_k)q\right|. \qquad (2.3.11)$$

Introducând notația $q_k = \sum_{i=c_{k-1}+1}^{c_k} p_i$, avem semnificația că pe următoarele m_k niveluri de gri din histograma transformată (adică pe nivelurile $z_{k-1}+1,\ldots,z_k = z_{k-1}+m_k$) trebuie plasați cei q_k pixeli care corespund nivelurilor de gri $c_{k-1}+1,\ldots,c_k$ din histograma inițială. Plasarea acestor q_k pixeli se va realiza pe un unic nivel de gri al histogramei transformate, și anume $y_k = \left[\frac{z_{k-1}+z_k}{2}\right]$.

Încheierea algoritmului

Procesul iterativ prezentat anterior se încheie după *Etapa* μ, când este îndeplinită condiția $q_1+\ldots+q_\mu = N$.

În histograma transformată sunt utilizate nivelurile de gri $y_1,\ldots,y_\mu$ pe care se plasează, respectiv, $q_1,\ldots,q_\mu$ pixeli.

Scanned with OKEN Scanner

Pixelii aparţinând nivelurilor de gri puţin populate în histograma iniţială ($p_i < q$) se cumulează pe grupuri cât mai apropiate de valoarea q şi sunt plasaţi pe un singur nivel de gri în histograma transformată; acesta constituie cazul $m_k = 1$ din descrierea generală din Etapa k. Pixelii aparţinând unui nivel de gri abundent populat în histograma iniţială ($p_i > q$) sunt preluaţi în totalitate şi plasaţi în mijlocul unui interval cu nuanţe de gri din histograma transformată; acesta constituie cazul $m_k > 1$ din descrierea generală din Etapa k.

Exemplul 2.3.6

Se consideră drept imagine iniţială teaca Hertwig în dezvoltarea mugurelui dentar din fig. 2.3.1, cu histograma din fig. 2.3.2, având numărul total de pixeli N = 796x1200 = 955200. Prin operaţia de egalizare a histogramei se redistribuie aceşti N pixeli pe toate cele n = 256 niveluri de gri, obţinându-se imaginea din fig. 2.3.11, cu histograma în fig. 2.3.12. Histograma transformată ocupă întregul domeniu al valorilor de gri, numărul ideal de pixeli ce ar fi trebuit plasat pe fiecare nivel fiind, conform algoritmului anterior prezentat, q = [955200/256] = [3731.25] = 3731 pixeli. Cum în histograma iniţială pe majoritatea nivelurilor de gri populate (adică între 56 şi 146) există mai mult de q pixeli, acest număr de pixeli se va transfera în histograma finală, plasându-se însă pe alte niveluri de gri decât în histograma iniţială şi corespunzând unor intervale de două sau mai multe niveluri de gri. De exemplu, numărul maxim de pixeli 17592 din histograma iniţială se regăseşte în histograma transformată în mijlocul unui interval ce conţine 5 niveluri de gri (deoarece valoarea 17592/3731 = 4.72 se aproximează, cu eroare minimă faţă o valoare întreagă, prin numărul 5).

Problema se pune similar şi pentru nivelurile de gri vecine celui cu număr maxim de pixeli, ceea ce înseamnă că în histograma transformată nivelurile de gri corespunzătoare acestei zone se populează la o distanţă de 4-5 niveluri (a se vedea reprezentarea grafică din fig. 2.3.12, pentru nivelurile de gri cuprinse între 150 şi 200, din care numai 11 sunt populate). Mai mult chiar, efectul de spaţiere în popularea nivelurilor de gri din histograma transformată este vizibil în tot domeniul [10, 250], fiecare din nivelurile populate conţinând un număr de pixeli apropiat de un multiplu de q de forma mq, $m \geq 3$. În schimb, primele câteva niveluri de gri populate din histograma iniţială (adică începând cu 56) conţin mai puţin de q pixeli, care se cumulează în histograma transformată pe grupuri de aproximativ q pixeli, plasându-se pe primele câteva niveluri de gri consecutive ($y_1 = 0, y_2 = 1, \ldots$).

Similar, ultimele câteva niveluri de gri populate din histograma iniţială (adică imediat înainte de 146) conţin mai puţin de q pixeli, care se cumulează în histograma transformată pe grupuri de aproximativ q pixeli, plasându-se pe ultimele câteva niveluri de gri consecutive ($y_{\mu-2} = 253$, $y_{\mu-1} = 254$, $y_\mu = 255$). De fapt, efectul cumulativ al procedeului de egalizare a histogramei devine evident prin compararea extremităţilor histogramei din fig. 2.3.12 cu extremităţile histogramei din fig. 2.3.9, obţinută printr-o simplă expandare liniară (care nu modifică numărul de pixeli de pe nivelurile de gri ale histogramei iniţiale din fig. 2.3.2, ci doar îi plasează pe alte niveluri, cu o spaţiere mai mare, pentru a ocupa întregul domeniu [0, 255]).

Scanned with OKEN Scanner

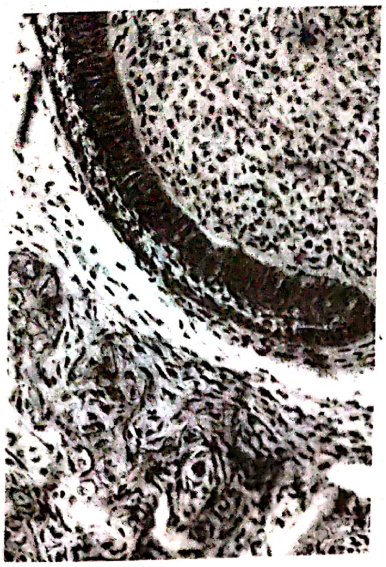
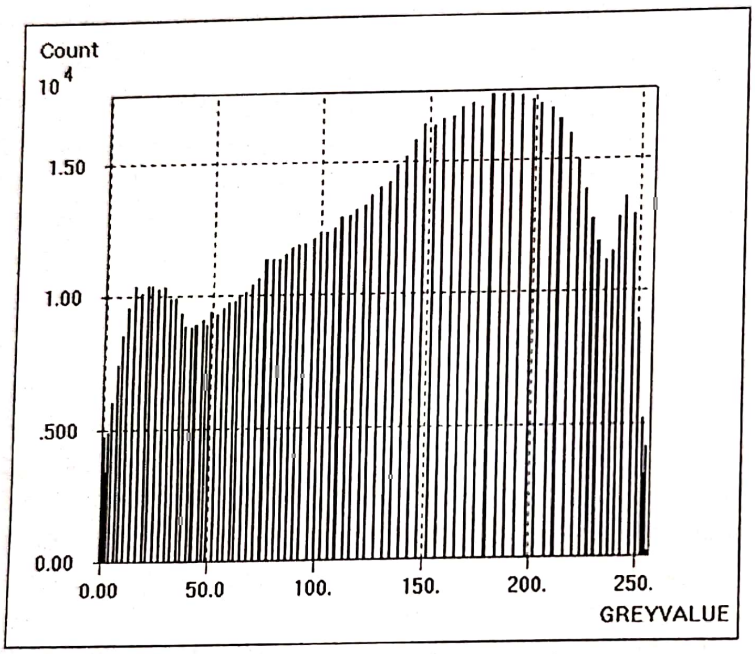


Fig. 2.3.11. Imagine rezultată ca urmare a îmbunătățirii contrastului, datorită redistribuirii nivelurilor de gri prin egalizarea histogramei.



Statistics	
file	7
parameter	GREYVALUE
counts	955200
lower bound	0
upper bound	255
underflow	0
in range	955200
overflow	0
classes	256
modul	1
min.count	0
max.count	17592
Statistics	
minimum	0
maximum	255
sum	1.21407e+008
medianvalue	127.502
mean	127.102
variance	5425.71
std.dev.	73.6594
skewness	-0.00112819
kurtosis	-1.20126

Fig. 2.3.12. Histograma de luminozitate corespunzătoare imaginii din fig. 2.3.11 şi tabelul cu informații statistice aferente.

Observație. În final, atragem atenția asupra faptului că manipularea contrastului trebuie disociată de modificarea rezoluției imaginii. Considerăm importantă această precizare deoarece, uneori, îmbunătățirea contrastului poate să creeze observatorului uman impresia unei rezoluții mai bune.

2.3.3. ELIMINAREA ZGOMOTELOR ŞI UMBRELOR

Imaginile digitale reale sunt afectate de o serie de imperfecțiuni sau perturbații, rezultate din modalitatea de achiziție. Aceste imperfecțiuni sunt de obicei clasificate în două categorii, şi anume: **zgomote** şi **umbre**. **Zgomotul** (*eng.* noise) se defineşte pentru o mulțime redusă de

Scanned with OKEN Scanner

pixeli învecinaţi (regiune mică), caracterizată printr-o variaţie în nuanţa de gri (respectiv de culoare), într-o zonă percepută ca având colorit uniform în imaginea originală (analogică). Termenul de "zgomot" este utilizat în limbajul specific prelucrării imaginilor prin analogie cu noţiunea tradiţională de zgomot, care afectează acurateţea unui semnal sonor. **Umbra** (*eng.* shadow) se defineşte pentru o mulţime numeroasă de pixeli învecinaţi (regiune mare), caracterizată prin nuanţe de gri mai închise comparativ cu ansamblul imaginii, şi este datorată, în principal, unei iluminări neuniforme a imaginii originale (analogică).

Tehnicile de procesare digitală a imaginilor permit eliminarea în mare măsură a imperfecţiunilor de tip zgomot sau umbre, prin **operaţii de filtrare** (*eng.* filtering). Aceste operaţii au drept rezultat modificarea nivelurilor de gri (culorilor) ale anumitor pixeli, strategia de modificare a fiecărui pixel luând în considerare vecinătatea respectivului pixel. Cu alte cuvinte, operaţiile de filtrare nu vizează pixelii ca entităţi individuale, ci ca elemente aparţinând anumitor mulţimi. În acest context, precizăm că în urma unei operaţii de filtrare nu toţi pixelii care au avut o aceeaşi nuanţă de gri (culoare) în imaginea iniţială sunt modificaţi la fel, spre deosebire de operaţiile asupra histogramei (a se vedea ajustarea luminozităţii şi manipularea contrastului). Această tratare selectivă este tocmai consecinţa modului de analiză ce se raportează la informaţiile aferente măştilor.

- **ELIMINAREA ZGOMOTELOR**

Zgomotul poate apărea pe parcursul achiziţiei, transmisiei sau procesării imaginii şi poate fi dependent sau nu de conţinutul imaginii. Din punct de vedere matematic, zgomotele pot fi descrise ca variabile aleatoare, ale căror caracteristici se definesc în conformitate cu tipul zgomotului. O aprofundare în această direcţie nu este relevantă din punctul de vedere al analizei de imagine asistată de calculator în morfologia microscopică.

Sub raportul nuanţelor de gri, pixelii afectaţi de zgomot apar adesea extrem de diferiţi faţă de pixelii din vecinătatea lor, neafectaţi de zgomot. Zgomotul nu poate fi eliminat complet, dar operaţiile de filtrare îl pot reduce în mod substanţial, datorită efectului de **netezire** (*eng.* smooth) a variaţiilor mari în nuanţele de gri sau culoare.

Operaţiile de filtrare fac apel la diverse tipuri de algoritmi. Cei mai frecvent utilizaţi sunt **algoritmii de mediere** şi **algoritmii gaussieni**. În cele ce urmează, vom detalia numai **algoritmii de mediere**, deoarece aceştia au o utilizare mult mai frecventă în prelucrarea imaginilor de morfologie microscopică. Principiul medierii se bazează pe aplicarea, pentru fiecare pixel din imaginea originală, a unei măşti sau nucleu, de dimensiuni 3x3, 5x5, 7x7, ..., LxL pixeli (L de valoare impară, vezi fig. 2.2.6, subsecţiunea 2.2.4), de aşa manieră încât pixelul investigat să se plaseze în centrul de simetrie al măştii. Nuanţa de gri (sau culoarea) ce se va atribui pixelului investigat se obţine ca o medie a valorilor pixelilor învecinaţi, selectaţi cu ajutorul măştii. În funcţie de tipul de medie calculată, algoritmii de filtrare se pot clasifica în: **filtrare medie** sau **"trece jos"** (*eng.* low pass filtering) şi **filtrare mediană** (*eng.* median filtering).

În **filtrarea medie** sau "trece jos", valoarea medie calculată pe mască este dată de media aritmetică.

Valoarea de gri $g(i,j)$ a pixelului din imaginea filtrată, ce corespunde pixelului cu aceeaşi poziţie (i,j) în imaginea iniţială, şi având nivelul de gri $f(i,j)$, se determină după formula:

$$g(i,j)=\frac{1}{w}\sum_{(m,n)\in M} f(m,n). \qquad (2.3.12)$$

În această formulă, w reprezintă numărul de pixeli care se găsesc în masca M aplicată pixelului cu poziția (i,j). Media pe masca M poate include sau nu pixelul curent (i,j) localizat în centrul de simetrie al măștii. Mărimea măștii trebuie să fie corelată cu conținutul imaginii. Pentru măşti prea mari, fenomenul de netezire are drept efect benefic îndepărtarea zgomotelor, dar este de obicei însoțit, ca efect negativ, de o anumită "înceţoşare" a imaginii, constând în deteriorarea contrastului şi a contururilor existente în imaginea inițială.

În **filtrarea mediană**, valoarea medie calculată pe mască este dată de mediana setului ordonat de niveluri de gri (culori). Într-o primă etapă, pixelii din masca M, de dimensiuni LxL, aplicată pixelului curent, sunt selectați din cadrul imaginii inițiale. În a doua etapă, pixelii din masca M sunt ordonați (crescător sau descrescător) după valoarea nivelului de gri. Se selectează valoarea de gri poziționată în mijlocul acestui şir ordonat, de lungime L^2. Valoarea de gri a pixelului curent $g(i,j)$ este înlocuită în imaginea filtrată cu valoarea mediană astfel calculată. O îmbunătățire a filtrării mediane se poate obține ca generalizare a procedeului descris. După ce pixelii din mască sunt ordonați în conformitate cu valoarea nivelului de gri, se exclud primii şi ultimii k pixeli ($k < L^2/2$). În loc de a selecta o singură valoare, prin considerarea medianei, se calculează media aritmetică a celor L^2-$2k$ niveluri de gri cuprinse între k+1 şi L^2-(k+1).

În general se consideră că algoritmii de filtrare de tip median produc o "înceţoşare" minimă, motiv pentru care sunt preferați pentru imaginile de microscopie.

Exemplul 2.3.7

Se consideră imaginea din fig. 2.3.13 reprezentând o zonă de pulpă roşie din splină, cordoanele Billroth alternând cu sinusoidele splenice. În contextul unei măsurători orientate pe dimensiunile cordoanelor Billroth, populația celulară din sinusoide poate fi interpretată drept zgomot. Pentru îndepărtarea acestui zgomot, pe imaginea inițială din fig. 2.3.13 se aplică un filtru median cu mască de 5x5 pixeli, cu rezultat în imaginea din fig. 2.3.14. Se observă şi apariția unui efect redus de înceţoşare, datorat filtrării de tip median.

Fig. 2.3.13. Splină, pulpă roşie – imagine inițială pentru procedeul de îndepărtare a zgomotului.

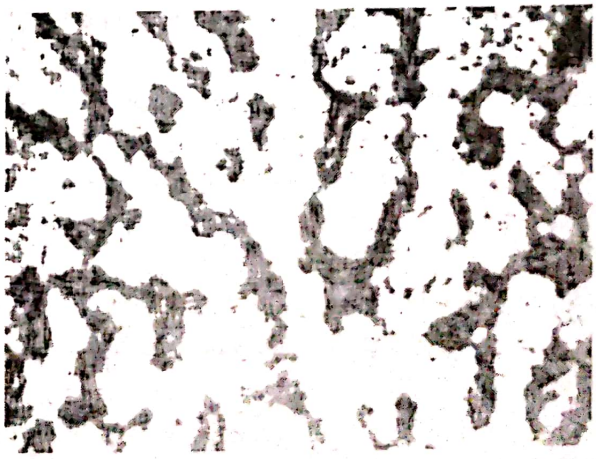


Fig. 2.3.14. Imagine rezultată în urma aplicării filtrului median pentru îndepărtarea zgomotului, cu mască 5x5 pixeli.

Scanned with OKEN Scanner

- **ELIMINAREA UMBRELOR**

Cele două tipuri de **filtrări** prezentate anterior, ca instrumente pentru eliminarea zgomotelor, pot fi utilizate şi pentru **eliminarea umbrelor**, considerând însă măşti de netezire de dimensiuni mari. În aceste condiţii de exploatare, drept rezultat al filtrării unei imagini cu iluminare neuniformă, se va obţine o imagine fără delimitarea şi claritatea structurilor, dar care pune în evidenţă tocmai neuniformitatea iluminării. Astfel, noua imagine conţine zone largi de gri deschis (corespunzând iluminării puternice) şi, respectiv, zone largi de gri închis (corespunzând iluminării slabe).

Printr-o succesiune de **operaţii aritmetice** adecvate (care folosesc atât imaginea iniţială, cât şi cea rezultată prin filtrare), precum şi prin eventuale **manipulări de histograme**, se poate ajunge la o imagine relativ uniformă ca şi iluminare, din care umbrele sunt îndepărtate. O tehnică conducând la acest efect constă în următorii paşi: crearea negativului pentru imaginea filtrată; medierea acestuia cu imaginea iniţială; deplasarea sau expandarea histogramei pentru imaginea rezultată din mediere.

Exemplul 2.3.8

Se consideră imaginea din fig. 2.3.15 (zonă de medulară renală) care prezintă o iluminare neuniformă. Pe această imagine se aplică de 4 ori consecutiv o operaţie de filtrare medie cu o mască foarte largă (49x49), obţinându-se imaginea din fig. 2.3.16. Pornind de la aceste două imagini se realizează imaginea din fig. 2.3.17, din care au fost eliminate umbrele, astfel încât prezintă un aspect uniform sub raportul iluminării. Utilizând aceeaşi imagine iniţială (fig. 2.3.15) şi aplicând o operaţie de filtrare mediană cu acelaşi tip de mască (ce conduce la imaginea din fig. 2.3.18) se poate ajunge la un rezultat similar privind îndepărtarea umbrelor (imaginea din fig. 2.3.19).

Fig. 2.3.15. Medulară renală – imagine iniţială neuniform luminată pentru procedeul de eliminare a umbrelor.

Scanned with OKEN Scanner

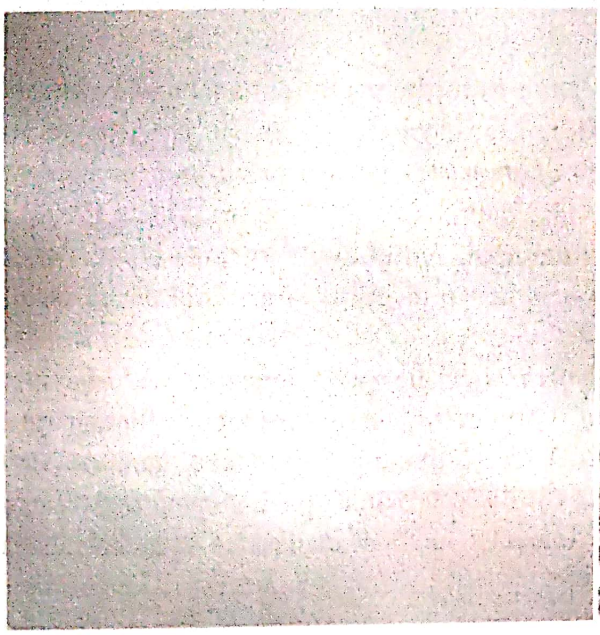


Fig. 2.3.16. Rezultatul filtrării medii realizată pe imaginea din fig. 2.3.15.

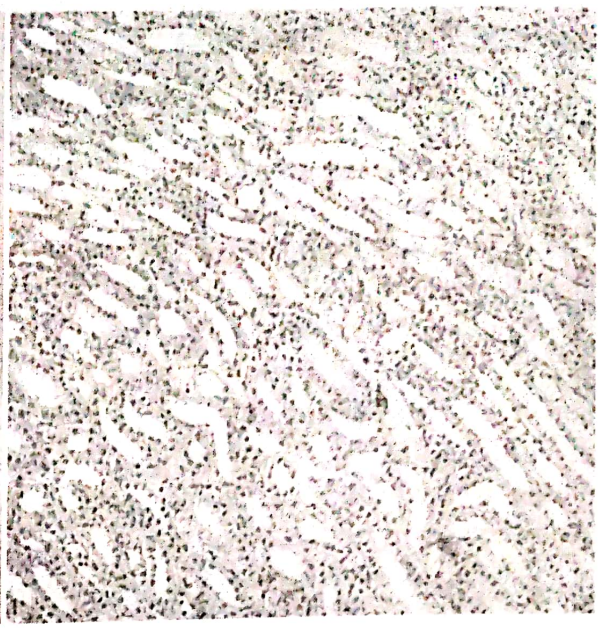


Fig. 2.3.17. Imagine cu luminozitate uniformă, rezultată din aplicarea filtrării medii pe imaginea inițială.

Fig. 2.3.18. Rezultatul filtrării mediane realizată pe imaginea din fig. 2.3.15.

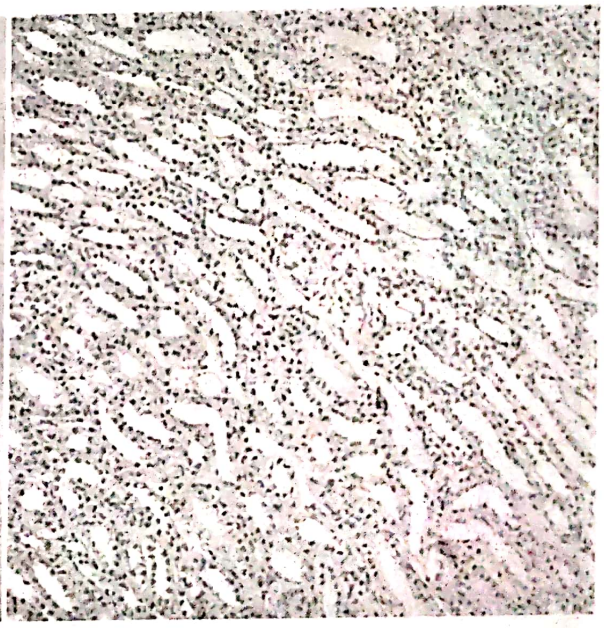


Fig. 2.3.19. Imagine cu luminozitate uniformă, rezultată din aplicarea filtrării mediane pe imaginea inițială.

2.3.4. DETECȚIA MUCHIILOR ȘI IDENTIFICAREA CONTURURILOR. SEPARAREA REGIUNILOR

O **muchie** (*eng.* edge) se definește ca zona între două regiuni adiacente, care prezintă niveluri de gri substanțial diferite. Așadar, detecția muchiilor dintr-o imagine presupune determinarea acelor zone în care funcția nivel de gri prezintă variații bruște. În imaginile de microscopie

optică specifice morfologiei normale şi patologice, detecţia muchiilor poate fi o problemă destul de dificilă, deoarece diferenţele dintre regiunile de interes sunt relativ mici, datorită elementelor histoarhitectonice (celularitate densă, stratificare, joncţiuni strânse), precum şi particularităţilor coloraţiilor tipice utilizate în realizarea preparatelor.

Identificarea punctelor ce aparţin unei muchii se bazează pe procedee matematice capabile să detecteze diferenţele importante între nivelurile de gri corespunzătoare unor pixeli învecinaţi. Aceste procedee utilizează **operatori (funcţii) diferenţiali bidimensionali** care extind, într-un spaţiu cu două dimensiuni, de natură discretă (specifică imaginilor digitale), conceptul de derivată a unei funcţii de o singură variabilă, cunoscut din analiza matematică clasică. De aceea, principiul respectivelor procedee poate fi expus într-o manieră simplă, considerând valorile luminozităţii (nivelul de gri) pixelilor dintr-o muchie, situaţi de-a lungul unei singure direcţii în planul imaginii. Astfel, luminozitatea poate fi privită ca o **funcţie de o singură variabilă**, a cărei reprezentare grafică arată de obicei ca în fig. 2.3.20.

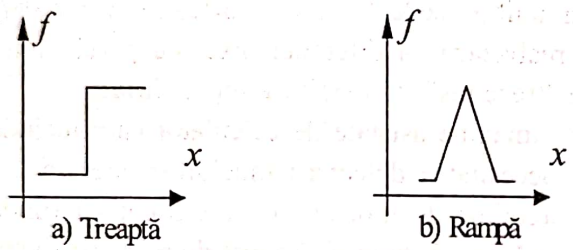


Fig. 2.3.20. Reprezentare grafică pentru modificarea nivelurilor de gri corespunzând unei muchii plasate pe verticală: (a) modificare de tip treaptă; (b) modificare de tip rampă.

Ambele grafice din fig. 2.3.20 corespund unei modificări în nivelurile de gri, de la o regiune mai întunecată (la stânga) la o regiune mai luminoasă (la dreapta). În fig. 2.3.20(a) modificarea se realizează printr-un salt (treaptă), iar în fig. 2.3.20(b) – printr-o rampă. Menţionăm că modelul de tip rampă este reprezentativ deoarece, în general, muchiile apar uşor neclare în imaginile digitale, ca rezultat al procesului de digitizare.

În timp ce o variaţie de tip treaptă este uşor de detectat, detecţia unei variaţii de tip rampă face neapărat apel la utilizarea diferenţelor de ordinul întâi şi doi, definite pentru un punct de abscisă i astfel:

- diferenţă de ordinul întâi (echivalent discret al derivatei de ordin întâi):
$$\nabla^1 f(i) = f(i+1) - f(i) \; ; \qquad (2.3.13a)$$
- diferenţă de ordinul doi (echivalent discret al derivatei de ordin doi):
$$\nabla^2 f(i) = \nabla^1(\nabla^1 f(i)) = \nabla^1 f(i+1) - \nabla^1 f(i) = f(i+2) - 2f(i+1) - f(i) \; . \qquad (2.3.13b)$$

Diferenţa (derivata) de ordinul întâi este zero în ambele regiuni cu nivel de gri constant (porţiunile orizontale de la stânga şi de la dreapta rampei) şi ia o valoare constantă pe parcursul tranziţiei de nivel de gri (pe rampa propriu-zisă). Diferenţa (derivata) de ordinul doi este zero în toate punctele (atât pe porţiunile orizontale, cât şi pe rampă), excepţie făcând începutul şi sfârşitul tranziţiei de nivel de gri. Pentru punctul de la începutul rampei (în vecinătatea regiunii întunecate), diferenţa (derivata) de ordinul doi este pozitivă, iar pentru punctul de la sfârşitul rampei (în vecinătatea regiunii luminoase), diferenţa (derivata) de ordinul doi este negativă. Astfel, devine evident faptul că nenulitatea diferenţei (derivatei) de ordinul întâi poate fi utilizată pentru a detecta prezenţa unei muchii, în timp ce semnul diferenţei (derivatei) de ordinul doi precizează învecinarea cu o zonă întunecată sau una luminoasă.

Scanned with OKEN Scanner

În situația ilustrativă discutată, funcția nivel de gri a fost considerată ca depinzând de un singur parametru, respectiv poziția pixelilor în raport cu o singură direcție. În cazul unei imagini, discuția trebuie purtată în raport cu două direcții (de exemplu direcțiile axelor de coordonate), motiv pentru care **diferența (derivata) de ordin întâi** se generalizează prin noțiunea de **operator de tip gradient**, iar **diferența (derivata) de ordin doi** prin noțiunea de **operator Laplace** (Laplacean). Prin exploatarea acestora în diferite maniere au fost dezvoltate mai multe tipuri de operatori destinați detecției muchiilor și, implicit, **tehnici** performante ce poartă același nume cu a operatorilor utilizați, dintre care amintim pe cele mai cunoscute: Sobel, Roberts, Prewitt, Kirsch, Laplace. Întrucât aplicarea acestora are drept efect o tratare neuniformă a pixelilor cu aceeași valoare a nivelurilor de gri, tehnicile amintite pot fi privite ca **metode speciale de filtrare**. În contextul prezentei lucrări, considerăm că discutarea suportului teoretic ce permite implementarea acestor tehnici ar depăși nivelul de formalizare matematică propus, motiv pentru care nu furnizăm alte detalii. Cititorul interesat poate consulta lucrările monografice recomandate în debutul capitolului.

Tot în cadrul detecției muchiilor trebuie amintit procedeul bazat pe **filtrarea "trece sus"** (*eng.* high-pass filtering). Această operație se realizează ca diferența pixel cu pixel dintre imaginea inițială, netezită prin filtrare medie sau "trece jos", și imaginea inițială însăși.

În problematica generală a prelucrării de imagine asistate de calculator care include procesări de imagini spațiale (în profunzime), necesitatea detectării muchiilor intervine în numeroase aplicații. În cadrul particular al imaginilor de morfologie microscopică (unde obiectivul principal îl constituie extragerea de trăsături geometrice pentru diverse structuri), detecția muchiilor este utilizată îndeosebi pentru determinarea automată a **contururilor** (*eng.* contour). În limbajul tehnic specific domeniului, relația dintre noțiunile de **muchie** și **contur** reiese din modul de definire a conturului, drept muchii conectate ce caracterizează forma unui obiect. Astfel, în cazul imaginilor morfologice dintr-un câmp microscopic, contururile unor structuri sau regiuni închise, cu o anumită semnificație biologică, se identifică în urma detectării muchiilor aferente.

În acest context, algoritmii de detectare a muchiilor prezentați anterior pot fi utilizați pentru extragerea contururilor. Având în vedere comentariile de la începutul subsecțiunii curente, asupra specificului imaginilor morfologice, precizăm că în extragerea contururilor ne putem confrunta cu următoarele două tipuri de dificultăți: (*i*) existența unor discontinuități relativ mici în muchiile ce aparțin contururilor; (*ii*) variații reduse în nuanțele de gri (coloritul) ale zonelor unde, în realitate, ar trebui detectate muchii. Eliminarea acestor neajunsuri necesită un tratament special și se bazează, în general, pe următoarele principii: în cazul (*i*), se folosesc tehnici care permit conectarea discontinuităților din muchii (de exemplu transformata Hough); în cazul (*ii*), se utilizează filtre de tip "albie" (*eng.* watershed), care, prin trasarea unor curbe de nivel (similare celor de pe hărțile geografice), ajută la separarea obiectelor ce se ating.

Exemplul 2.3.9

Se consideră imaginea inițială (foliculi tiroidieni) din fig. 2.3.21, asupra căreia se realizează o operație de filtrare cu operator de tip Sobel. Rezultatul apare în fig. 2.3.22, în care detecția muchiilor permite identificarea cu ușurință a contururilor. Pixelii ce aparțin muchiilor și, implicit, contururilor, apar în nuanțe de gri foarte luminos. Zonele cu colorație uniformă în imaginea inițială apar în nuanțe de gri foarte întunecat.

Scanned with OKEN Scanner

În acest sens, se va remarca faptul că, în fiecare folicul tiroidian, coloidul (colorat inițial în gri închis), cât și spațiul artefactual adiacent (colorat inițial în alb) capătă nuanțe de gri foarte întunecat în imaginea finală, ele fiind separate de pixelii foarte luminoși corespunzători conturului (ca tranziție de culoare).

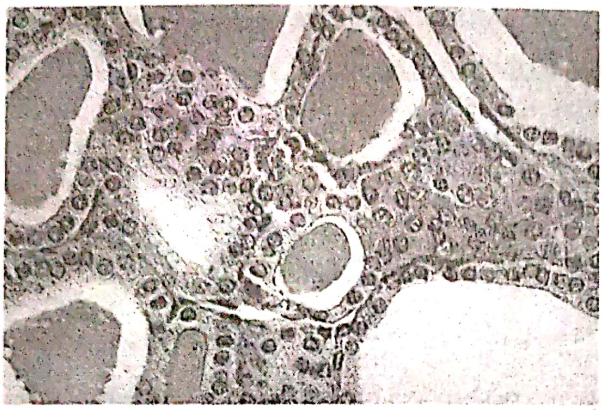


Fig. 2.3.21. Foliculi tiroidieni – imagine inițială pentru procedeul de detecție a muchiilor.

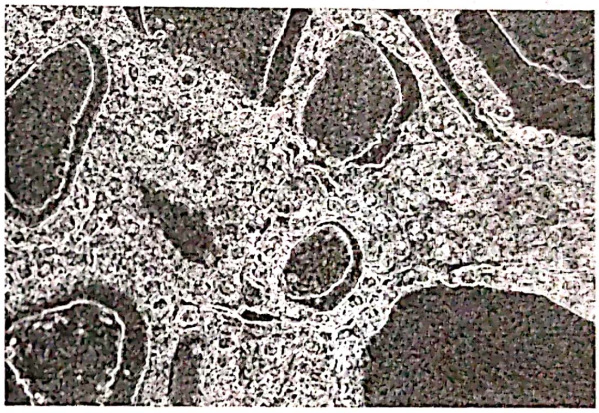


Fig. 2.3.22. Rezultatul procedeului de detecție a muchiilor prin filtrare cu operator de tip Sobel.

În prelucrarea unor imagini morfologice există situații când prezintă interes **delimitarea** (*eng.* delineate) **a două regiuni**, fără a pune în evidență efectiv muchia. O atare abordare se poate realiza cu ajutorul unui algoritm care, în zona de tranziție a nivelurilor de gri dintre două regiuni, modifică valorile nivelurilor de gri ale pixelilor de așa manieră încât toți pixelii din această zonă să poată fi atribuiți (pe baza nivelurilor de gri rezultate) la una din cele două regiuni adiacente. Considerând o mască M de tipul celor prezentate în subsecțiunea 2.2.4, fig. 2.2.6, să notăm cu x, respectiv y, valoarea de gri a pixelului central înainte și, respectiv, după modificarea nivelului de gri. De asemenea, vom folosi notația generică z pentru nivelul de gri al unui pixel din zona acoperită de mască (evident, mai puțin pixelul central) în imaginea inițială. Pentru a decide dacă un pixel aparține zonei de tranziție a nivelurilor de gri și, în consecință, dacă asupra sa trebuie operată o modificare de luminozitate, se utilizează un prag P. Cu aceste precizări, principiul algoritmului se poate formula astfel:

$$y=\begin{cases} x, & \text{pentru } (\max\limits_{M} z-\min\limits_{M} z)\le P, \\ \min\limits_{M} z, & \text{pentru } (\max\limits_{M} z-\min\limits_{M} z)>P \ \&\ (\max\limits_{M} z-x)\ge(x-\min\limits_{M} z), \\ \max\limits_{M} z, & \text{pentru } (\max\limits_{M} z-\min\limits_{M} z)>P \ \&\ (x-\min\limits_{M} z)>(\max\limits_{M} z-x). \end{cases} \qquad (2.3.14)$$

Exemplul 2.3.10

Se consideră aceeaşi imagine iniţială (foliculi tiroidieni) din fig. 2.3.21, asupra căreia se realizează o operaţie de delimitare a regiunilor, conform formulei (2.3.14), cu masca $M = 25x25$ şi pragul $P = 127$. Rezultatul apare în fig. 2.3.23, în care se constată efectul repartizării pixelilor din zona de tranziţie a nivelurilor de gri la una din cele două regiuni adiacente, caracterizate prin colorații ferme. În fig. 2.3.23 se constată absenţa conturului propriu-zis pentru coloid (existent în fig. 2.3.22), dar coloidul este bine delimitat de spaţiul artefactual adiacent, pe baza diferenţei de culoare între cele două regiuni (spre deosebire de fig. 2.3.22).

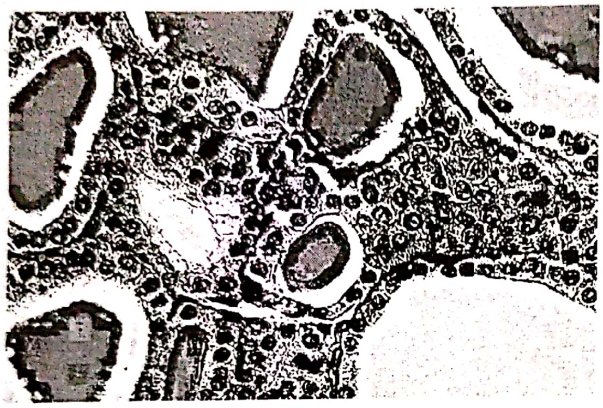


Fig. 2.3.23. Rezultatul procedeului de delimitare a regiunilor.

În anumite clasificări a operaţiilor ce pot fi aplicate imaginilor digitale, extragerea contururilor şi delimitarea regiunilor sunt grupate sub denumirea de operaţii de **clarificare** (*eng.* sharpening).

2.4. SEGMENTAREA IMAGINILOR DIGITALE

Termenul de **segmentare** (*eng.* segmentation) a imaginii se referă la partiţionarea unei imagini într-o mulţime de regiuni care o acoperă. În majoritatea situaţiilor, regiunile reprezintă arii din imagine, cu o anumită semnificaţie. Regiunile pot fi definite drept grupuri de pixeli având atât un contur, cât şi o formă particulară (de exemplu cerc, elipsă sau diverse poligoane). Pe de altă parte, regiunile pot fi caracterizate de un anumit grad de omogenitate, în raport cu o anumită proprietate: nivel de gri (sau culoare), modulul gradientului funcţiei nivel de gri etc.

Segmentarea poate avea două obiective, şi anume: (*i*) descompunerea imaginii în părţi, pentru o analiză ulterioară; (*ii*) efectuarea unei modificări de reprezentare. Ca o clasificare generală a metodelor de segmentare, se poate vorbi despre segmentare **bazată pe regiuni** şi, respectiv, **bazată pe curbe**. De asemenea, tot din perspectiva unei clasificări, trebuie subliniat faptul că există metode generale de tip "de jos în sus" (*eng.* bottom-up), care nu folosesc cunoştinţe apriorice referitoare la structurile din imagine, precum şi modalităţi de segmentare care utilizează informaţii specifice domeniului. În primul caz, o imagine constă dintr-un număr de structuri situate pe un fond uniform, faţă de care există o diferenţă mare de contrast, iar procesul segmentării poate fi dus la bun sfârşit folosind tehnici de procesare de nivel scăzut. În al doilea caz, procesul de segmentare trebuie să facă apel la o serie de date suplimentare şi să coopereze cu alte module de nivel înalt, utilizate ulterior segmentării în procesarea imaginii.

Scanned with OKEN Scanner

2.4.1. BINARIZAREA DE TIP PRAG

Binarizarea imaginii reprezintă o metodă de segmentare adecvată separării entităților dintr-o imagine de microscopie morfologică, în vederea extragerii trăsăturilor acestora. Algoritmii de binarizare pot fi aplicați cu succes în cazul în care structurile din imagine au un nivel de gri relativ constant (caracteristici de culoare similare), ce contrastează cu nivelul de gri asociat fondului, diferența fiind suficient de netă. Aceasta face ca separarea regiunilor din imagine să poată fi realizată fără prea mari dificultăți prin impunerea unui prag de binarizare, pe baza căruia fiecare pixel din imagine va fi clasificat ca fiind pixel **de fond** sau pixel **obiect**. Se spune că un atare procedeu de binarizare este de **tip prag** (*eng.* thresholding). Folosind notația generică y pentru un pixel al imaginii de destinație (binarizate), valoarea binară a lui y va rezulta din nivelul de gri notat generic x, din imaginea inițială, conform următoarei reguli:

$$y = \begin{cases} 1\,(\text{alb}), & \text{pentru } x > P, \\ 0\,(\text{negru}), & \text{pentru } x \le P, \end{cases} \tag{2.4.1}$$

unde P desemnează pragul de decizie pentru separarea structurilor de fundal.

Astfel, prin binarizarea de tip prag, o imagine digitală cu scală de gri este transformată într-o imagine digitală binară sau echivalent alb-negru (în care fiecare din pixeli poate lua numai una din două valori, de exemplu 0 pentru negru și 1 pentru alb). În cazul în care fundalul imaginii este întunecat (nuanțe de gri închis), imaginea binară rezultată în urma segmentării poate fi interpretată drept conținând în planul apropiat pixelii ce aparțin structurilor de interes (cu alb), iar în alt plan îndepărtat, sub formă de fundal, pixelii nesemnificativi (cu negru).

Conform relației (2.4.1), operația de binarizare are un caracter local. Noua valoare asociată pixelului curent supus analizei depinde numai de valoarea pragului și de valoarea nivelului de gri propriu; alți factori, cum ar fi poziția sa, sau valorile nivelului de gri dintr-o vecinătate a sa, nu sunt luați în considerație. O astfel de schemă de binarizare este avantajoasă din punctul de vedere al efortului de calcul, pentru o imagine de dimensiune NxN numărul de operații necesare (adică complexitatea algoritmului) fiind de ordinul valorii N^2. Drept consecință, algoritmul poate fi ușor implementat și testat.

Alegerea unui **prag de binarizare** P corect este esențială pentru a obține rezultate în concordanță cu formele din imagine. Acest prag poate fi ales interactiv sau poate fi determinat utilizând diferite proceduri de tipul celor discutate mai jos. Variații relativ mici ale valorii pragului de binarizare P pot influența semnificativ rezultatul segmentării. De exemplu, în cazul imaginilor cu fundal gri închis, pot să apară următoarele situații generatoare de erori: (*i*) o valoare a lui P prea mică poate conduce la înglobarea în cadrul structurilor a multor pixeli ce aparțin, de drept, fundalului; (*ii*) o valoare a lui P prea mare poate conduce la omiterea unor părți din zonele ce aparțin, de drept, structurilor.

Ca fundament teoretic, funcționalitatea binarizării de tip prag rezultă din **proprietățile histogramei**. Pentru imagini conținând structuri cu niveluri de gri relativ uniforme și diferite mult față de fondul imaginii, histograma are un caracter **bimodal**, în sensul că este formată din două grupări (moduri) relativ bine delimitate. Astfel, dacă pixelii fondului au niveluri de gri închis, ei determină apariția unei grupări în stânga histogramei, în timp ce pixelii structurilor, presupuși a avea niveluri de gri deschis, conduc la apariția unei grupări în dreapta histogramei. În general, cele două grupări nu sunt complet separate, între vârfurile grupărilor existând populații de pixeli care ocupă nivelurile de gri intermediare. Alegerea unui nivel de

Scanned with OKEN Scanner

prag corespunzător valorii de minim aflate între cele două maxime ale histogramei ar putea constitui un criteriu acceptabil din punct de vedere intuitiv.

Totuşi, utilizarea soluţiei bazate pe valoarea de minim în aplicaţiile practice nu este întotdeauna uşor de aplicat, deoarece în multe cazuri condiţia de alură bimodală a histogramei nu este îndeplinită. Mai mult, chiar în cazul în care această condiţie este îndeplinită, pot să apară inconveniente de factura celor enumerate mai jos:

- Dacă zona cuprinsă între cele două maxime este foarte întinsă, nu este posibilă localizarea precisă a minimului;
- În situaţia în care sunt prezente mai multe puncte de minim local, nu există un criteriu fundamentat ştiinţific pentru alegerea celui mai bun dintre acestea;
- Pentru o imagine afectată de zgomot, determinarea unei valori cu semnificaţie reală de minim pe histograma ideală (din care ar lipsi zgomotele) este incertă.

Datorită problemelor multiple pe care le poate ridica segmentarea prin binarizare, platformele software destinate procesării numerice a imaginilor pun la dispoziţia utilizatorului atât variante **interactive**, cât şi **automate**, pentru exploatarea histogramei în scopul binarizării. De obicei, procedurile de binarizare automată au la bază **algoritmi iterativi de determinare a valorii de prag** P. Prezentăm mai jos un astfel de algoritm, cu o structură suficient de simplă:

- *etapa 1*: se alege o valoare iniţială pentru pragul P; valorile uzuale ce pot fi folosite sunt media nivelului de gri din imagine sau $(x_{max}+x_{min})/2$, unde x_{max}, x_{min} notează valoarea maximă şi, respectiv, minimă a nivelului de gri;
- *etapa 2*: se binarizează imaginea iniţială pe baza nivelului de prag ales;
- *etapa 3*: se calculează mediile M_1, M_2 pentru fiecare dintre grupările separate de pragul P;
- *etapa 4*: se calculează o nouă valoare de prag $P = (M_1+M_2)/2$;
- *etapa 5*: se repetă paşii 2-4 până la realizarea convergenţei şirului de valori determinate pentru pragul P.

Exploatarea histogramei pentru determinarea automată a pragului de binarizare se poate face şi în manieră **probabilistică**, considerând nivelurile de gri din imagine drept valori pe care le ia o **variabilă aleatoare discretă**. Aceste valori se grupează în două clase, pentru care se determină mediile şi varianţele. Fără a intra în detalii, precizăm că principiul algoritmului constă în a modifica iterativ cele două clase în sensul minimizării sumei varianţelor, ponderate cu mediile claselor aferente.

Deoarece este convenabil a colora în alb regiunile ce se doresc a fi evidenţiate în urma binarizării şi, respectiv, în negru restul imaginii (cu semnificaţie de fundal), în locul algoritmului de binarizare tip prag poate fi folosit un **algoritm de binarizare de tip interval**. Folosind notaţiile generice x şi y cu aceeaşi semnificaţie ca la binarizarea tip prag, şi desemnând prin L, H, limita inferioară şi, respectiv, superioară a intervalului, algoritmul de binarizare de tip interval se defineşte (pentru 256 niveluri de gri în imaginea iniţială) astfel:

$$y=\begin{cases}0\,(\text{negru}), & \text{pentru } 0\le x<L,\\ 1\,(\text{alb}), & \text{pentru } L\le x\le H,\\ 0\,(\text{negru}), & \text{pentru } H<x\le 255.\end{cases} \tag{2.4.2}$$

Ca şi în situaţia comentariilor referitoare la binarizarea tip prag, mici variaţii în alegerea intervalului $[L, H]$ (adică lungimea intervalului, cât şi modul de plasare în domeniul $[0, 255]$) pot conduce la modificări semnificative în selectarea regiunilor de interes, prin procesul de binarizare.

Scanned with OKEN Scanner

Exemplul 2.4.1

Se consideră imaginea din fig. 2.4.1, reprezentând tubuli dentinari în secțiune longitudinală într-un preparat șlif, cu histograma de luminozitate din fig. 2.4.2. Se urmărește ca printr-o operație de segmentare prin binarizare să evidențiem regiunile corespunzătoare structurilor tubulare, colorate, în fig. 2.4.1, în nuanțe de gri închis. Se constată că histograma nu îndeplinește condiția de bimodalitate, ea posedând un singur maxim, în vecinătatea nivelului de gri 170. Ilustrăm operația de binarizare în următoarele trei situații care utilizează intervale diferite pentru gama de nuanțe de gri, considerate semnificative pentru regiunile ce trebuie puse în evidență în fig. 2.4.1. În toate cele trei situații, în imaginile binare, funcționează următorul tip de corespondență: zonele albe trebuie privite ca asociate tubulilor dentinari, iar fundalul negru trebuie înțeles drept dentină intertubulară.

1. Pentru valorile de gri din intervalul [0, 50], conform formulei (2.4.2) cu $L = 0$ și $H = 50$, rezultă imaginea binară din fig. 2.4.3.

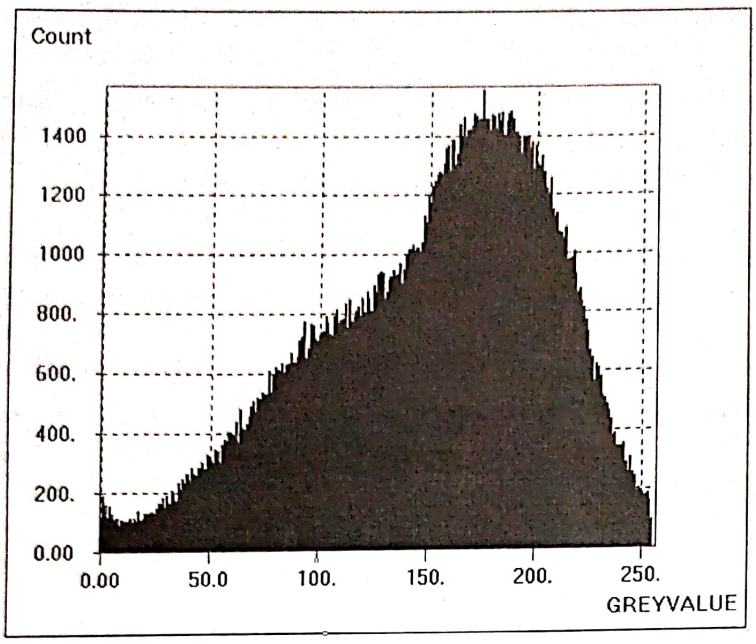
2. Pentru valorile de gri din intervalul [0, 110], conform formulei (2.4.2) cu $L = 0$ și $H = 110$, rezultă imaginea binară din fig. 2.4.4.

3. Pentru valorile de gri din intervalul [0, 150], conform formulei (2.4.2) cu $L = 0$ și $H = 150$, rezultă imaginea binară din fig. 2.4.5.

Imaginea inițială a fost anume aleasă numai cu structuri longitudinale pentru ca efectele binarizării în condiții variate să devină ușor vizibile, prin diferențele în grosime a zonelor albe, corespunzătoare tubulilor. În acest sens, este sugestivă o examinare comparativă a fig. 2.4.3, fig. 2.4.4 și fig. 2.4.5, care demonstrează că, odată cu creșterea intervalului de nuanțe de gri aplicabil în binarizare (drept corespondent al regiunilor de interes), crește grosimea regiunilor de interes. Subliniem faptul că situația 2 oferă cea mai bună reprezentare binară a tubulilor dentinari, în raport cu imaginea inițială. În situația 1, imaginea binarizată permite recunoașterea a mai puțini tubuli dentinari, comparativ cu imaginea inițială. Din contră, în situația 3, dimensiunile tubulilor dentinari sunt mai mari decât cele reale, zonele de alb corespunzătoare acestora incluzând, în mod fals, și teritorii din dentina peritubulară.

Fig. 2.4.1. Dinte șlif – tubuli dentinari în secțiune longitudinală – imagine inițială pentru procedeul de segmentare tip prag.

Scanned with OKEN Scanner



Statistics	
file	2
parameter	GREYVALUE
counts	189000
lower bound	0
upper bound	255
underflow	0
in range	189000
overflow	0
classes	256
modul	1
min.count	8
max.count	1567
Statistics	
minimum	0
maximum	255
sum	2.85227e+007
medianvalue	160.322
mean	150.914
variance	2951.43
std.dev.	54.3271
skewness	-0.514627
kurtosis	-0.401044

Fig. 2.4.2. Histograma de luminozitate corespunzătoare imaginii din fig. 2.4.1 şi tabelul cu informaţii statistice aferente.

Fig. 2.4.3. Imagine binară rezultată în urma segmentării, utilizând domeniul de niveluri de gri [0, 50].

Fig. 2.4.4. Imagine binară rezultată în urma segmentării, utilizând domeniul de niveluri de gri [0, 110].

Scanned with OKEN Scanner

Fig. 2.4.5. Imagine binară rezultată în urma segmentării, utilizând domeniul de niveluri de gri [0, 150].

Observaţie. Binarizarea funcţiei de nivel de gri constituie cea mai frecvent utilizată posibilitate de segmentare a unei imagini. Algoritmii de binarizare pot fi însă aplicaţi în acelaşi mod şi pentru cazul când, în locul funcţiei ce exprimă nivelul de gri al pixelilor, se utilizează o funcţie construită pe baza operatorilor de tip gradient, discutaţi însubsecţiunea 2.3.4, la detecţia muchiilor şi identificarea contururilor.

2.4.2. BINARIZAREA PRIN CREŞTEREA REGIUNILOR

Funcţia nivel de gri a unei imagini poate fi întrebuinţată, fără a face apel la histogramă, în cadrul tehnicii denumite "**creşterea regiunilor**" (*eng.* region growing). Această tehnică se bazează pe **conectivitatea** pixelilor cu niveluri de gri apropiate, permiţând ca, prin startarea de la un pixel cu o anumită valoare de gri, să se realizeze expandarea la o regiune conţinând numai nuanţele de gri dintr-un interval anume definit (căruia, evident, îi aparţine nuanţa de gri a punctului de start). Pentru stabilirea conectivităţii pixelilor, se pot folosi ambele tipuri de vecinătăţi, V_4 şi V_8. În implementările software, intervalul nuanţelor apropiate de gri pe baza cărora are loc creşterea regiunii poate fi controlat **interactiv** de către utilizator sau poate fi setat **automat**, ca rezultat al unei caracterizări statistice a coloritului întregii imagini.

Exemplul 2.4.2

Se consideră imaginea din fig. 2.4.6 (1200 x 868 pixeli), reprezentând o etapă evolutivă a mugurelui dentar, cu dentina în curs de formare. Ilustrăm operaţia de creştere a regiunilor, realizată interactiv, cu acelaşi tip de conectivitate V_4, pentru următoarele trei situaţii:

1. Pornind din pixelul de coordonate (0, 732), cu valoarea nivelului de gri 52 şi permiţând creşterea în domeniul de niveluri de gri [0, 65], se obţine regiunea albă din fig. 2.4.7 (corespunzătoare benzii de dentină depuse).
2. Pornind din pixelul de coordonate (0, 854), cu valoarea nivelului de gri 220 şi permiţând creşterea în domeniul de niveluri de gri [170, 255], se obţine regiunea albă din fig. 2.4.8 (corespunzătoare zonei subjacente dentinei, cu prelungiri odontoblastice, şi unor spaţii artefactuale interodontoblastice).
3. Pornind din pixelul de coordonate (2, 844), cu valoarea nivelului de gri 148 şi permiţând creşterea în domeniul de niveluri de gri [66, 169], se obţine regiunea albă din fig. 2.4.9 (corespunzătoare pulpei dentare cu odontoblaste periferice şi, parţial, celulelor din organul smalţului).

Scanned with OKEN Scanner

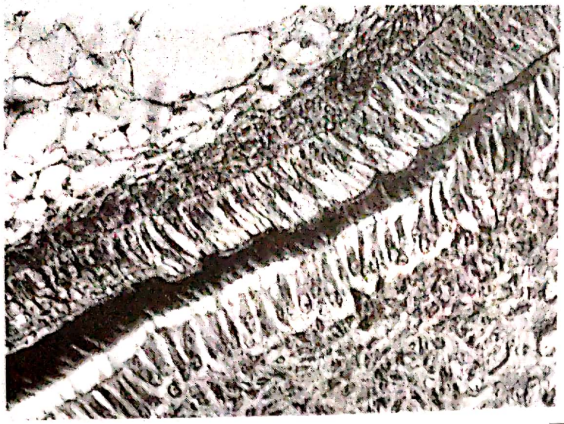


Fig. 2.4.6. Mugure dentar – dentină în curs de formare – imagine inițială pentru procedeul de segmentare prin creșterea regiunilor.

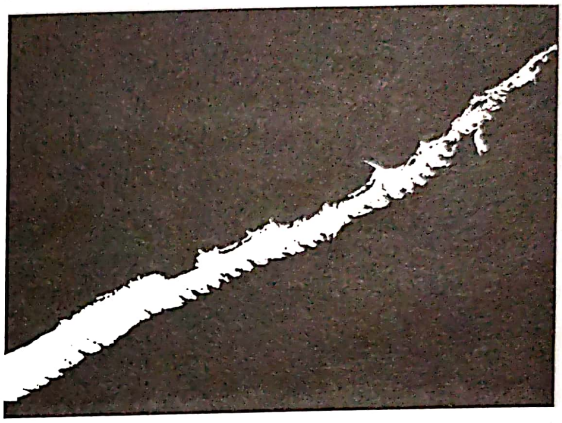


Fig. 2.4.7. Imagine binară rezultată în urma segmentării prin creșterea regiunilor (interval [0, 65]), pornind din pixelul cu coordonatele (0, 732), având valoarea nivelului de gri 52.

Fig. 2.4.8. Imagine binară rezultată în urma segmentării prin creșterea regiunilor (interval [170, 255]), pornind din pixelul cu coordonatele (0, 854), având valoarea nivelului de gri 220.

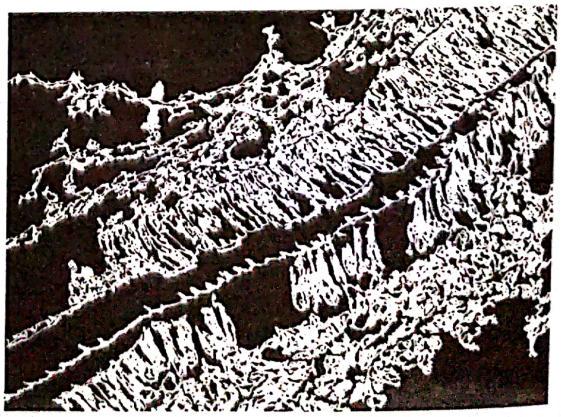


Fig. 2.4.9. Imagine binară rezultată în urma segmentării prin creșterea regiunilor (interval [66, 169]), pornind din pixelul cu coordonatele (2, 844), având valoarea nivelului de gri 148.

Scanned with OKEN Scanner

2.5. OPERAŢII ASUPRA IMAGINILOR BINARE

Imaginile binare (*eng.* binary images), construite pentru a pune în evidenţă regiuni de interes ce vor face obiectul determinărilor cantitative, conţin frecvent o serie de imperfecţiuni (rezultate din procedeul de binarizare), care pot afecta semnificativ efectuarea măsurătorilor. Din această cauză, înainte de efectuarea măsurătorilor, trebuie operate corecţii ale imperfecţiunilor existente, astfel încât regiunile albe din imaginea binară să se apropie cât mai mult, ca dimensiune şi formă, de regiunile colorate corespondente din imaginea iniţială. Corecţiile se pot realiza prin **operaţii de morfologie binară** şi **operaţii booleene**. Prima categorie presupune acţiuni localizate la nivelul regiunilor albe dintr-o imagine, în timp ce a doua categorie desemnează acţiuni ce vizează ansamblul imaginii.

2.5.1. MORFOLOGIE BINARĂ

Morfologia binară (*eng.* binary morphology) acţionează la nivelul structurilor existente în imaginea binară, având drept efect modificări din alb în negru sau invers a anumitor pixeli, în baza unor criterii bine definite. Principalele operaţiile prin care se controlează aceste modificări sunt **dilatarea**, **erodarea**, **deschiderea** şi **închiderea** imaginilor binare, la care se adaugă **eliminarea unor regiuni** şi **umplerea găurilor** din imaginile binare.

- **DILATAREA IMAGINILOR BINARE**

Dilatarea (*eng.* dilation) este o operaţie morfologică aplicată imaginilor binare, prin care, după anumite principii, se adaugă zonelor colorate în alb, noi pixeli albi. Cu alte cuvinte, prin operaţia de dilatare se schimbă din negru în alb culoarea unor pixeli învecinaţi (în baza unei măşti M, de tipul celor prezentate în subsecţiunea 2.2.4) zonelor de alb; înainte de aplicarea operaţiei pixelii respectivi fac parte din fundalul întunecat, iar după efectuarea operaţiei ei aparţin regiunilor luminoase.

Descrierea algoritmică a operaţiei de dilatare se bazează pe noţiunea de **translare** a unei mulţimi de pixeli X printr-un vector de poziţie t, în urma căreia rezultă mulţimea de pixeli translată X_t, definită astfel:

$$X_t = \{x+t \mid x \in X\}. \tag{2.5.1}$$

Cu aceasta, operaţia de dilatare a unei imagini binare B cu ajutorul măştii M, notată $B \oplus M$, este definită prin formula:

$$B \oplus M = \bigcup_{b \in B} M_b, \tag{2.5.2}$$

unde M_b înseamnă masca M aplicată unui pixel arbitrar b, din imaginea binară B.

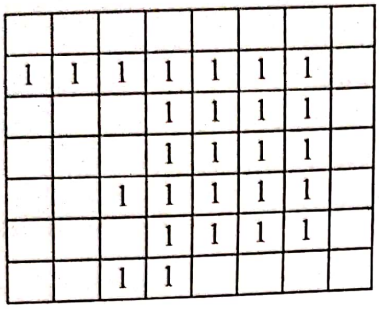
Acţiunea operaţiei de dilatare cu masca 3x3 pixeli din fig. 2.5.1(b), asupra imaginii binare din fig. 2.5.1(a), este prezentată în fig. 2.5.2.

Operaţia de dilatare se poate aplica iterativ, numărul iteraţiilor fiind corelat cu numărul de pixeli suplimentari ce se doresc a fi ataşaţi imaginii, ţinând cont de efectele tipice detaliate în cele ce urmează.

Efectele notabile ale operaţiei de dilatare constau în îngroşarea structurilor în raport cu imaginea martor iniţială (nedilatată), netezirea marginilor şi umplerea găurilor mici, izolate înăuntrul structurilor. Din această cauză, în imaginile de microscopie optică, şi, în particular, pentru celule, dilatarea se aplică cu scopul ca imaginile rezultate să semene mai bine cu structurile reale. Prin prisma acestor observaţii, menţionăm faptul că aplicarea unei măşti de

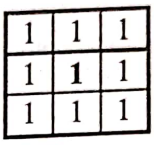
Scanned with OKEN Scanner

dimensiuni mici, consecutiv, de un număr mare de ori, are aproximativ același efect ca și aplicarea unei măști de dimensiuni mari, consecutiv, de un număr redus de ori. Mai mult, atragem atenția asupra faptului că un număr exagerat de mare de iterații în aplicarea operației de dilatare va conduce la o imagine cu toți pixelii albi.



(a)

Fig. 2.5.1(a). Reprezentare convențională a unei imagini binare.



(b)

Fig. 2.5.1(b). Reprezentare convențională a unei măști de 3x3 pixeli.

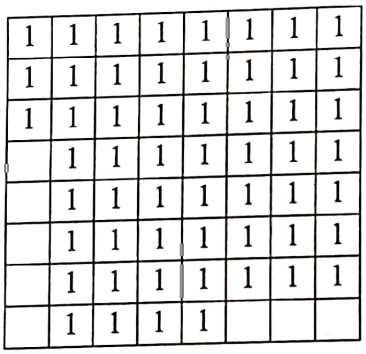


Fig. 2.5.2. Rezultatul operației de dilatare pe imaginea din fig. 2.5.1(a), cu masca din fig. 2.5.1(b).

În contextul comentariilor de mai sus, precizăm că unele medii software pun la dispoziție o procedură specială de dilatare, denumită **dilatare maximă** (*eng.* ultimate dilation), care, indiferent de tipul de mască folosit și numărul de iterații aplicate, împiedică fuzionarea regiunilor albe, lăsând întotdeauna o linie subțire de fundal (negru) între ele.

- **ERODAREA IMAGINILOR BINARE**

Erodarea (*eng.* erosion) este o operație morfologică aplicată imaginilor binare, prin care, după anumite principii, se elimină, din zonele colorate în alb, pixeli albi. Cu alte cuvinte, prin operația de erodare se schimbă din alb în negru culoarea unor pixeli învecinați (în baza unei măști M, de tipul celor prezentate în subsecțiunea 2.2.4) zonelor de alb; înainte de aplicarea operației, pixelii respectivi fac parte din regiunile luminoase, iar după efectuarea operației ei aparțin fundalului întunecat.

Operația de erodare a unei imagini binare B cu ajutorul măștii M, notată $B \ominus M$, este definită prin formula:

$$B \ominus M = \{b \mid b + m \in B \; \forall m \in M\}, \tag{2.5.3}$$

în sensul că, înaintea aplicării operației, imaginea binară B era alcătuită din pixelii notați generic $b + m$, din care vor dispare cei notați m, corespunzători măştii M.

Acțiunea operației de erodare cu masca 3x3 pixeli din fig. 2.5.1(b), asupra imaginii binare din fig. 2.5.1(a), este prezentată în fig. 2.5.3.

Operația de erodare se poate aplica iterativ, numărul iterațiilor fiind corelat cu numărului de pixeli suplimentari ce se doresc a fi eliminați din imagine, ținând cont de efectele tipice detaliate în cele ce urmează.

Efectele notabile ale operației de erodare constau în subțierea structurilor în raport cu imaginea martor inițială (neerodată) şi îndepărtarea pixelilor sau grupurilor mici de pixeli izolați din imagine (în majoritate legați de zgomotul alb-negru, denumit "sare şi piper"). Ca şi în cazul dilatării, aplicarea unei măşti de dimensiuni mici, consecutiv, de un număr mare de ori, are aproximativ acelaşi efect ca şi aplicarea unei măşti de dimensiuni mari, consecutiv, de un număr redus de ori. Mai mult, atragem atenția asupra faptului că un număr exagerat de mare de iterații în aplicarea operației de erodare va conduce la o imagine cu toți pixelii negri.

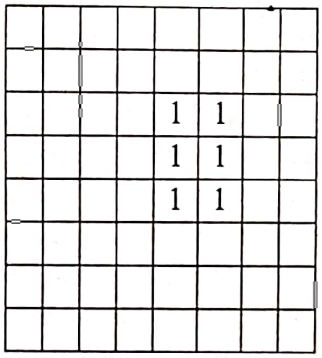


Fig. 2.5.3. Rezultatul operației de erodare pe imaginea din fig. 2.5.1(a), cu masca din fig. 2.5.1(b).

În contextul comentariilor de mai sus, precizăm că unele medii software pun la dispoziție o procedură specială de erodare, denumită **erodare maximă** (*eng.* ultimate erosion), care, indiferent de tipul de mască folosit şi numărul de iterații aplicate, împiedică ştergerea completă a regiunilor albe.

- **ÎNCHIDEREA ŞI DESCHIDEREA IMAGINILOR BINARE**

Operația de erodare nu reprezintă în mod strict inversa operației de dilatare, motiv pentru care compunerea celor două operații prezintă interes.

Deschiderea (*eng.* opening) este o operație morfologică aplicată imaginilor binare, prin care se realizează o erodare urmată de o dilatare.

Operația de deschidere a unei imagini binare B cu ajutorul măştii M, notată $B \circ M$, este definită prin formula:

$$B \circ M = (B \ominus M) \oplus M, \tag{2.5.4}$$

unde operațiile $\ominus, \oplus$ au fost definite conform (2.5.3), respectiv (2.5.2).

Acțiunea operației de deschidere cu masca 3x3 pixeli din fig. 2.5.1(b), asupra imaginii binare din fig. 2.5.1(a), este prezentată în fig. 2.5.4.

Închiderea (*eng.* closing) este o operație morfologică aplicată imaginilor binare, prin care se realizează o dilatare urmată de o erodare.

Scanned with OKEN Scanner

Operația de închidere a unei imagini binare B cu ajutorul măștii M, notată $B \bullet M$, este definită prin formula:

$$B \bullet M = (B \oplus M) \Theta M, \tag{2.5.5}$$

unde operațiile $\oplus, \Theta$ au fost definite conform (2.5.2), respectiv (2.5.3).

Acțiunea operației de închidere cu masca 3x3 pixeli din fig. 2.5.1(b), asupra imaginii binare din fig. 2.5.1(a), este prezentată în fig. 2.5.5.

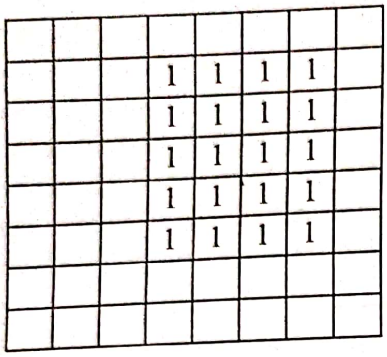


Fig. 2.5.4. Rezultatul operației de deschidere pe imaginea din fig. 2.5.1(a), cu masca din fig. 2.5.1(b).

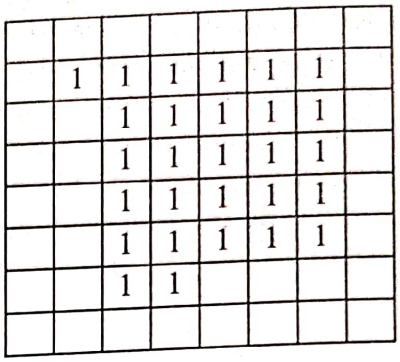


Fig. 2.5.5. Rezultatul operației de închidere pe imaginea din fig. 2.5.1(a), cu masca din fig. 2.5.1(b).

O comparație între fig. 2.5.1(a) (imaginea inițială) și fig. 2.5.4, respectiv fig. 2.5.5, pune în evidență faptul că atât deschiderea, cât și închiderea cu masca considerată lasă neschimbate zonele rectangulare masive de alb, deosebirile manifestându-se la nivelul pixelilor marginali.

- **ELIMINAREA UNOR REGIUNI DIN IMAGINILE BINARE**

Prin operația de **eliminare** (*eng.* scrap) se îndepărtează regiuni (grupuri de pixeli albi conectați), pe baza unor criterii formulate cu referire la ariile acestora. În general, formularea invocă un interval de tipul [aria_minimă, aria_maximă], iar regiunile eliminate pot fi selectate de așa manieră încât ariile lor să se situeze fie în interiorul intervalului respectiv, fie în afara lui. O mențiune aparte trebuie făcută cu privire la eliminarea grupurilor cu puțini pixeli, care au semnificație de zgomot, pentru care intervalul definit mai sus se alege cu aria_minimă egală cu zero, și aria_maximă egală cu o valoare stabilită în funcție de cunoștințele apriorice asupra elementelor utile din imagine.

- **UMPLEREA GĂURILOR ÎN IMAGINILE BINARE**

Operația de **umplere a găurilor** (*eng.* fill holes) permite umplerea oricărei găuri (mulțime cu pixeli negri) care există înăuntrul structurilor albe din imaginea binară. Fiind vorba despre o operație care se realizează automat, găurile se definesc strict ca arii din interiorul unor zone albe închise (în particular contururi). În cazul în care zona albă este fisurată (în particular conturul prezintă o discontinuitate), procedura automată nu mai identifică o gaură și, drept urmare, nu se va produce efectul de umplere. Astfel, pot să apară discordanțe între modul de finalizare a acestei operații morfologice efectuate automat, și rezultatul scontat de observatorul uman (care nu a perceput existența fisurii, respectiv a discontinuității în contur).

Scanned with OKEN Scanner

Exemplul 2.5.1

Se consideră imaginea din fig. 2.2.3 (reprezentând o secţiune şlif prin smalţ şi dentină, cu fisuri şi lamele la nivelul smalţului) pe care o reproducem în fig. 2.5.6, cu scopul de a facilita urmărirea rezultatelor operaţiilor morfologice, prin comparare cu imaginea iniţială din fig. 2.5.6. Prin operaţia de binarizare realizată cu intervalul de nuanţe de gri [0, 92], corespunzând regiunilor de interes, se obţine imaginea binară din fig. 2.5.7, în care nivelurile de gri cuprinse în intervalul [0, 92] sunt transformate în alb, iar celelalte în negru (conform formulei (2.4.2) cu $L = 0$ şi $H = 92$). Pentru a creşte acurateţea imaginii din fig. 2.5.7, se practică eliminarea regiunilor mici, cu arii cuprinse între 0-800 pixeli (interpretate ca zgomote), obţinându-se imaginea din fig. 2.5.8. Ulterior, imaginea din fig. 2.5.8 este supusă unei operaţii de dilatare cu mască 7x7, repetată de 3 ori, rezultând imaginea din fig. 2.5.9. Pe aceasta, se realizează o operaţie de erodare în aceleaşi condiţii (mască 7x7, repetată de 3 ori), obţinându-se imaginea din fig. 2.5.10. O comparaţie între fig. 2.5.8 şi fig. 2.5.9 arată creşterea zonelor de alb în urma operaţiei morfologice de dilatare. Prin aplicarea erodării asupra imaginii din fig. 2.5.9 se ajunge, în fig. 2.5.10, la o imagine binară foarte asemănătoare cu imaginea binară din fig. 2.5.8. Trebuie totuşi remarcate contururile mai netede din fig. 2.5.10, precum şi o mai mare compactitate a zonelor de alb. Tehnica dilatării, urmată de erodare, constituie o practică curentă pentru ameliorarea imaginilor rezultate din segmentare, în sensul apropierii formelor şi dimensiunilor zonelor albe de ceea ce reprezintă structurile de interes în imaginea originală.

Fig. 2.5.6. Ansamblu smalţ-dentină – imagine iniţială pentru aplicarea unor operaţii binare.

Fig. 2.5.7. Imagine rezultată în urma procedurii de segmentare, pe care se vor aplica operaţii de morfologie binară.

Scanned with OKEN Scanner

Fig. 2.5.8. Rezultatul aplicării procedurii de eliminare a regiunilor nesemnificative pe imaginea din fig. 2.5.7.

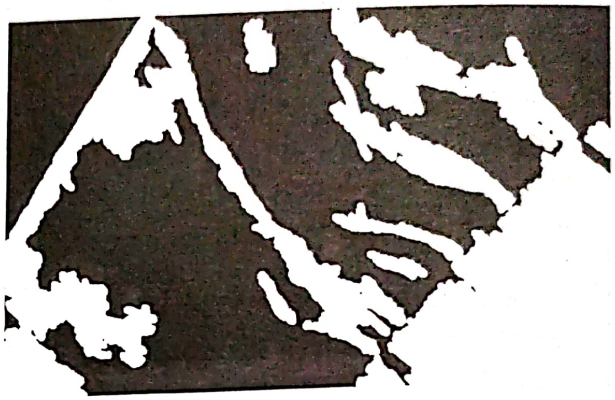


Fig. 2.5.9. Rezultatul aplicării operației morfologice binare de dilatare pe imaginea din fig. 2.5.8.

Fig. 2.5.10. Rezultatul aplicării operației morfologice binare de erodare pe imaginea din fig. 2.5.9.

Exemplul 2.5.2

Se consideră imaginea din fig. 2.5.11, reprezentând un canal interlobular al unei glande salivare, înconjurat de acini. Prin binarizarea realizată cu intervalul de nuanțe de gri [0, 100] alocat culorii albe (conform formulei (2.4.2), cu L = 0, H = 100) rezultă imaginea din fig. 2.5.12, unde structura de interes este canalul situat central. Lumenul acestui canal are drept corespondent, în imaginea binară din fig. 2.5.12, zona neagră din mijloc. Prin executarea unei operații de umplere a găurilor asupra imaginii binare din fig. 2.5.12 se obține imaginea binară din fig. 2.5.13. Se constată că operația a avut drept efect nu numai umplerea găurii mari, corespondente lumenului, ci s-a finalizat cu albirea tuturor zonelor din imaginea din fig. 2.5.12 – care corespund definiției tehnice de gaură într-o imagine binară.

Scanned with OKEN Scanner

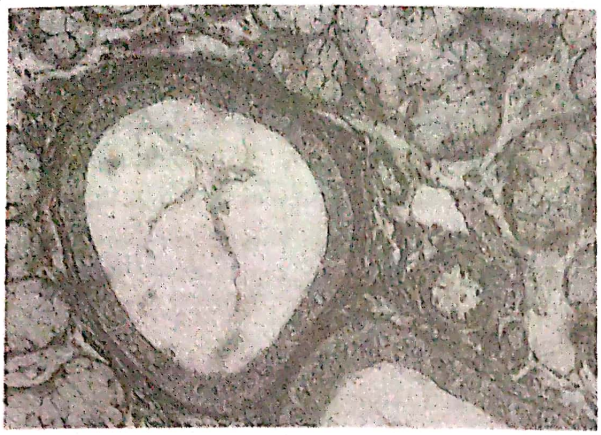


Fig. 2.5.11. Glandă salivară – canal interlobular – imagine inițială pentru procedeul de umplere a găurilor.

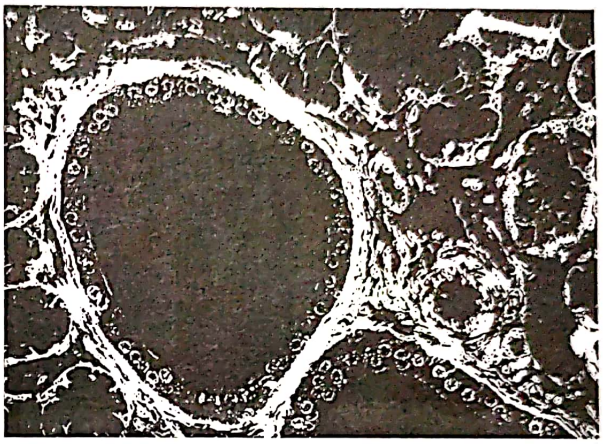


Fig. 2.5.12. Rezultatul aplicării operației de segmentare pe imaginea din fig. 2.5.11.

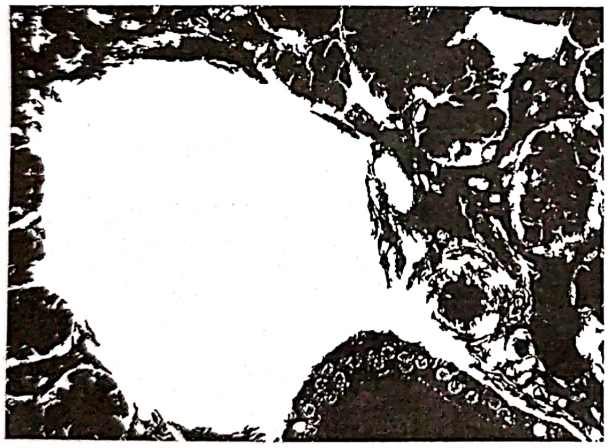


Fig. 2.5.13. Rezultatul operației de umplere a găurilor aplicată pe imaginea din fig. 2.5.12.

2.5.2. OPERAȚII BOOLEENE

Operațiile booleene (*eng.* boolean) se definesc la nivelul întregului ansamblu de pixeli ai imaginii binare, conform **operatorilor logici** ŞI (*eng.* AND), SAU (*eng.* OR), SAU-EXCLUSIV (*eng.* XOR), NU (*eng.* NOT). Acești operatori se aplică pixelilor cu aceleaşi coordonate (notate generic i, j), fiind definite respectiv prin tabelele din fig. 2.5.14(a), (b), (c), (d). Semnificația valorilor este 0 pentru negru şi 1 pentru alb. Primii trei operatori binari necesită doi operanzi: $x_1(i,j)$ (notația generică pentru valoarea unui pixel din prima imagine binară) şi $x_2(i,j)$ (notația generică pentru valoarea unui pixel din a doua imagine binară). Al patrulea operator este unar, necesitând un singur operand $x(i,j)$ (notația generică pentru valoarea unui pixel din imagine). Rezultatul acestor operații, $y(i,j)$, desemnează valoarea unui pixel generic din imaginea binară finală.

Scanned with OKEN Scanner

$y(i,j)$	$x_1(i,j)$	$x_2(i,j)$
0	0	0
0	0	1
0	1	0
1	1	1

(a)

(a) operatorul ŞI (AND);
$y(i,j)=x_1(i,j)\,\text{SI}\,x_2(i,j)$.

$y(i,j)$	$x_1(i,j)$	$x_2(i,j)$
0	0	0
1	0	1
1	1	0
1	1	1

(b)

(b) operatorul SAU (OR);
$y(i,j)=x_1(i,j)\,\text{SAU}\,x_2(i,j)$.

$y(i,j)$	$x_1(i,j)$	$x_2(i,j)$
0	0	0
1	0	1
1	1	0
0	1	1

(c)

(c) operatorul SAU-EXCLUSIV (XOR);
$y(i,j)=x_1(i,j)\,\text{SAU - EXCLUSIV}\,x_2(i,j)$.

$y(i,j)$	$x(i,j)$
1	0
0	1

(d)

(d) operatorul NU (NOT);
$y(i,j)=\text{NU}\,x(i,j)$.

Fig. 2.5.14. Definirea operațiilor logice.

Exemplul 2.5.3

Ne plasăm în contextul exemplului 2.5.1, considerând imaginea binară din fig. 2.5.7 rezultată prin operația de segmentare descrisă în cadrul exemplului respectiv. Acesteia i se aplică operatorul boolean NU (NOT), obținându-se imaginea binară din fig. 2.5.15 (în care zonele de alb din fig. 2.5.7 apar negre, iar zonele de negru din fig. 2.5.7 apar albe). Ca o completare la exemplul 2.5.1, merită precizat faptul că imaginea binară din fig. 2.5.15 ar putea fi obținută direct prin binarizarea imaginii monocrome din fig. 2.5.6, considerând drept regiuni de interes cele corespunzătoare intervalului de nuanțe de gri [93, 255] (cărora, astfel, li se alocă culoarea albă).

Fig. 2.5.15. Rezultatul aplicării operatorului boolean NU (NOT) pe imaginea din fig. 2.5.7.

Scanned with OKEN Scanner

Exemplul 2.5.4

Ne plasăm în contextul exemplului 2.4.2, considerând imaginile binare din fig. 2.4.7 şi fig. 2.4.8, rezultate prin operaţii de creştere a regiunilor în condiţiile discutate în acel exemplu. Aplicând operatorul boolean SAU (OR) acestor două imagini binare, se obţine imaginea binară din fig. 2.5.16, care reuneşte zonele albe din cele două imagini. Atragem atenţia asupra faptului că, pe acest principiu, pot fi reunite într-o singură imagine zonele albe corespunzătoare unor structuri de interes, identificate separat în mai multe imagini binare.

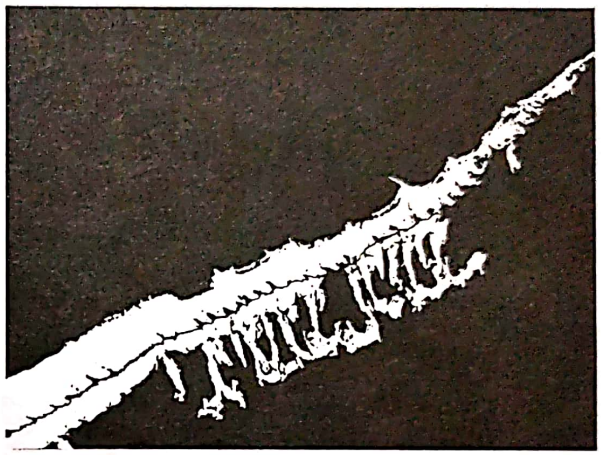


Fig. 2.5.16. Rezultatul aplicării operatorului boolean SAU (OR) pe imaginile din fig. 2.4.7 şi fig. 2.4.8.

Exemplul 2.5.5

Ne plasăm în contextul exemplului 2.5.2, considerând imaginile binare din fig. 2.5.12 şi, respectiv, fig. 2.5.13. Aplicând operatorul boolean SAU-EXCLUSIV (XOR) acestor două imagini binare, se obţine imaginea binară din fig. 2.5.17, în care au culoare albă numai acei pixeli ce corespund pixelilor cu valori diferite în fig. 2.5.12 şi fig. 2.5.13. Precizăm că pe acest principiu pot fi puse în evidenţă toate diferenţele ce apar între două imagini binare.

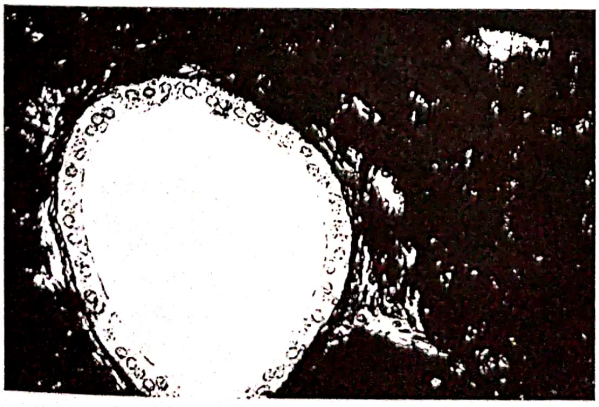


Fig. 2.5.17. Rezultatul aplicării operatorului boolean SAU EXCLUSIV (XOR) pe imaginile din fig. 2.5.12 şi fig. 2.5.13.

Scanned with OKEN Scanner

Exemplul 2.5.6

Ne plasăm tot în contextul exemplului 2.5.2, considerând imaginea binară din fig. 2.5.12. Pentru a selecta canalul central, se creează imaginea binară din fig. 2.5.18, care prezintă drept zonă albă un cerc cu diametrul corespunzător structurii de interes. Aplicând operatorul boolean ŞI (AND) imaginilor binare din fig. 2.5.12 şi fig. 2.5.18, rezultă imaginea din fig. 2.5.19. În fig. 2.5.19 au culoare albă numai acei pixeli ce corespund pixelilor cu valoare de alb atât în fig. 2.5.12, cât şi în fig. 2.5.18. Precizăm că pe acest principiu pot fi puse în evidenţă toate elementele comune a două imagini binare. Tehnica mascării prin operatorul ŞI (AND) cu o imagine binară conţinând o figură geometrică adecvată este utilizată curent pentru separarea unei regiuni de interes dintr-o imagine care conţine foarte multă informaţie.

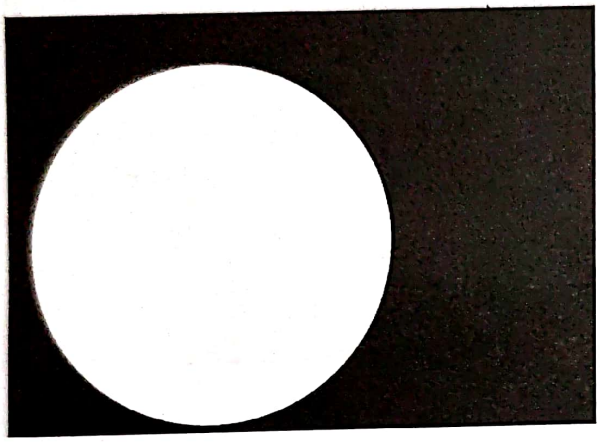


Fig. 2.5.18. Imagine binară creată în scopul selectării structurii de interes din fig. 2.5.12.

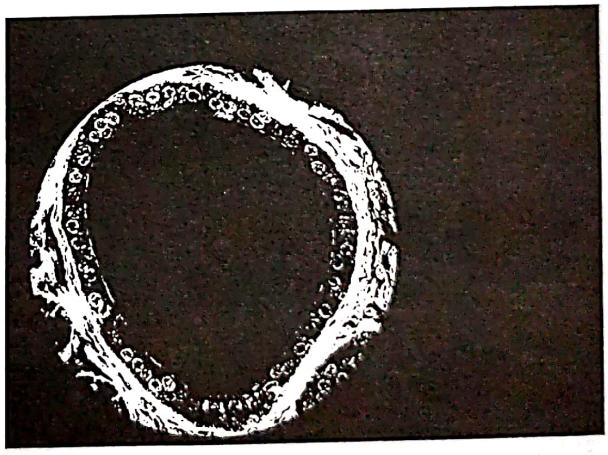


Fig. 2.5.19. Rezultatul aplicării operatorului boolean ŞI (AND) pe imaginile din fig. 2.5.12 şi fig. 2.5.18.

2.6. MĂSURARE / MĂSURĂTORI

Executarea **măsurătorilor** (*eng.* measurement) reprezintă pasul final în procesarea de imagine destinată morfologiei microscopice. Măsurătorile pot viza **proprietăţile geometrice** (care nu ţin cont de valoarea de culoare sau nuanţă de gri a pixelilor) şi **densitometrice/texturale** (care ţin cont de valoarea de culoare/nuanţă de gri a pixelilor).

În domeniul morfologiei normale şi patologice, interesul este focalizat în principal pe **trăsăturile geometrice** ale regiunilor. Ele pot viza următoarele categorii de parametri:

Scanned with OKEN Scanner

descriptori geometrici standard, descriptori bifazici specifici morfometriei celulare şi **descriptori utilizaţi pentru interpretări stereologice.**

Sub raportul implementărilor concrete, trebuie subliniat faptul că majoritatea pachetelor software orientate pe procesarea de imagine conţine funcţii predefinite pentru măsurarea descriptorilor geometrici standard, în timp ce *realizarea unor proceduri specifice de morfometrie celulară şi stereologie revine de obicei în sarcina utilizatorului.*

2.6.1. DESCRIPTORI GEOMETRICI STANDARD

Datorită complexităţii problemelor de morfometrie, produsele software existente la ora actuală pun la dispoziţie instrumente de calcul pentru un număr foarte mare de **parametri.** Dintre aceştia, cei mai frecvent utilizaţi sunt: aria, perimetrul, diametrul echivalent (prin raportare la un cerc), factori de formă (în raport cu circularitatea), axele elipsei echivalente (ce poate fi asociată regiunii, având acelaşi moment de inerţie cu al regiunii), orientarea semiaxelor pentru elipsa echivalentă.

În cele ce urmează, vom prezenta **principiile de calcul** limitându-ne doar la câţiva dintre aceşti descriptori, pentru care formulările matematice nu depăşesc gradul de complexitate propus în această lucrare. În acest scop, vom nota cu R mulţimea tuturor pixelilor din regiunea ale cărei trăsături geometrice se extrag, iar cu (i,j) coordonatele unui pixel generic din imagine.

Aria se calculează cu formula:

$$A = \sum_{(i,j)\in R} 1 , \tag{2.6.1}$$

ceea ce înseamnă numărarea tuturor pixelilor din regiune.

În scopul evaluării **perimetrului** unei regiuni, trebuie mai întâi determinaţi pixelii ce aparţin frontierei regiunii discutate. Pentru identificarea acestora, se testează vecinătăţile tuturor pixelilor din R. Pentru orice pixel (i,j) din interiorul lui R, vecinătatea considerată $V(i,j)$ va conţine numai pixeli din R, ceea ce în limbajul teoriei mulţimilor înseamnă $V(i,j)\setminus R=\Phi$, unde prin Φ se notează mulţimea vidă. În acelaşi timp, pentru orice pixel (i,j) de pe frontiera lui R, vecinătatea considerată $V(i,j)$ va conţine şi pixeli exteriori lui R, ceea ce, în limbajul teoriei mulţimilor, înseamnă $V(i,j)\setminus R\neq\Phi$. Astfel, frontiera este privită ca o mulţime F alcătuită din pixelii lui R care fac parte din a doua categorie discutată mai sus:

$$F=\{(i,j)\in R \mid V(i,j)-R\neq\Phi\}. \tag{2.6.2}$$

Să presupunem că frontiera conţine K pixeli. Pentru a evalua perimetrul, adică lungimea frontierei, ordonăm pixelii din mulţimea F într-o succesiune $(i_0,j_0),(i_1,j_1),\ldots,(i_{K-1},j_{K-1})$, pentru care perechile vecine înseamnă pixeli vecini, considerând şi vecinătatea ultimei perechi cu prima. În evaluarea perimetrului va trebui să ţinem cont de modul în care se învecinează doi pixeli; fiecare învecinare pe orizontală şi verticală va contribui la valoarea perimetrului cu câte o unitate, iar fiecare învecinare pe diagonală va contribui la valoarea perimetrului cu $\sqrt{2}$ unităţi. O atare diferenţiere se poate realiza testând modul de învecinare a pixelilor, ca fiind de tip V_4 (conectivitate pe laturi) sau de tip $V_8 \setminus V_4$ (conectivitate pe colţuri). Cu acestea, formula de calcul a perimetrului se poate exprima astfel:

Scanned with OKEN Scanner

$$P = \sum_{\substack{(i_{k+1},j_{k+1}) \in V_4(i_k,j_k) \\ k=0,\ldots,K;\, K+1 \overset{\Delta}{=} 0}} 1 + \sum_{\substack{(i_{k+},j_{k+1}) \in V_8(i_k,j_k) \setminus V_4(i_k,j_k) \\ k=0,\ldots,K;\, K+1 \overset{\Delta}{=} 0}} \sqrt{2} \,. \tag{2.6.3}$$

Având la dispoziție aria A, în ipoteza când regiunea se apropie ca formă de un disc (a se vedea factorii de formă vis-à-vis de circularitate), se poate calcula **diametrul echivalent** pentru regiune, cu formula:

$$D = 2\sqrt{\frac{A}{\pi}} \,. \tag{2.6.4}$$

Odată calculate aria A și perimetrul P, se poate determina o măsură a circularității regiunii R, care reprezintă un **factor de formă** în raport cu **circularitatea**, definit astfel:

$$C_1 = \frac{4\pi A}{P^2} \,. \tag{2.6.5}$$

Se constată că pentru o regiune perfect circulară $C_1 = 1$, astfel încât abaterile de la circularitate se pot cuantifica prin depărtarea de valoarea 1.

Un **factor de formă** cu funcționalitate mai largă poate fi definit fără a se face apel direct la aria A a regiunii, ci luând în considerare modul de plasare a celor K puncte ale frontierei în raport cu centrul de greutate al regiunii. Coordonatele $(\bar{i}, \bar{j})$ ale centrului de greutate fiind definite prin:

$$\bar{i} = \frac{1}{A} \sum_{(i,j) \in R} i \,, \quad \bar{j} = \frac{1}{A} \sum_{(i,j) \in R} j \,, \tag{2.6.6}$$

factorul de formă anunțat se exprimă astfel:

$$C_2 = \frac{\sigma_R}{m_R} \,, \tag{2.6.7}$$

unde

$$m_R = \frac{1}{K} \sum_{k=0}^{K-1} \| (i_k, j_k) - (\bar{i}, \bar{j}) \| \tag{2.6.8}$$

reprezintă media distanței de la centrul de greutate la frontieră, iar

$$\sigma_R = \left(\frac{1}{K} \sum_{k=0}^{K-1} [\| (i_k, j_k) - (\bar{i}, \bar{j}\} \| - m_R]^2 \right)^{1/2} \tag{2.6.9}$$

reprezintă deviația standard a distanței de la centrul de greutate la frontieră. Se constată că, pentru o regiune perfect circulară, $C_2 = 0$, creșterea abaterilor de la circularitate manifestându-se prin creșterea valorii lui C_2.

Pentru studierea principiilor de calcul aferente altor parametri geometrici, cititorul interesat poate consulta lucrările monografice recomandate în prezentarea generală a acestui capitol.

Exemplul 2.6.1

Se consideră imaginea din fig. 2.6.1, reprezentând pulpă dentară în zona centrală, intens celularizată. Prin aplicarea unei operații de segmentare, urmată de eliminarea regiunilor ce nu prezintă relevanță pentru măsurători (cu arii mai mici de 300 pixeli^2 și, respectiv,

Scanned with OKEN Scanner

mai mari de 3000 pixeli^2), rezultă imaginea din fig. 2.6.2, în care structurile de interes (celulele) apar albe. Pe această imagine se execută operații de măsurare, parametrii urmăriți fiind aria (conform formulei (2.6.1)) și perimetrul structurilor de interes (conform formulei (2.6.3)). Rezultatele măsurătorilor indică existența a 145 de regiuni, corespondența dintre arie și perimetru fiind următoarea:

- două celule cu aria minimă (303 pixeli^2) au perimetrul de 80.77 pixeli, respectiv 98.33 pixeli;
- celula cu perimetrul minim (75.84 pixeli) are o arie de 312 pixeli^2;
- celula cu aria maximă (2694 pixeli^2) are un perimetru de 695.55 pixeli, care este totodată și perimetrul maxim existent în secvența de măsurători.

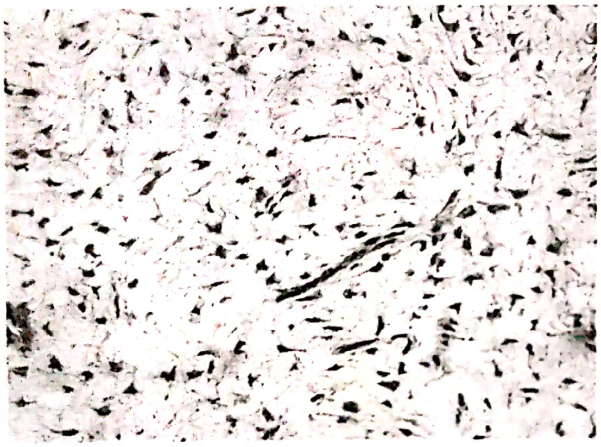


Fig. 2.6.1. Pulpă dentară – zonă centrală – imagine inițială pentru măsurători.

Fig. 2.6.2. Imagine rezultată prin aplicarea succesivă a operațiilor de segmentare și de eliminare a regiunilor considerate nesemnificative pe imaginea din fig. 2.6.1.

2.6.2. UNITĂȚI DE MĂSURĂ ȘI CONVERSII

Din modul de exprimare a formulelor de calcul din subsecțiunea precedentă, se observă că rezultatele măsurătorilor sunt furnizate în **pixeli** pentru caracteristici unidimensionale (lungimi), respectiv **pixeli^2** pentru cele bidimensionale (arii). În cazul când se impune o **conversie în unități metrice**, aceasta este realizată în baza **rezoluției** imaginii digitale, utilizată drept imagine de start în succesiunea de prelucrări numerice (cu condiția ca rezoluția să nu se modifice până în momentul efectuării măsurătorilor). Dacă rezoluția imaginii este n [ppi] (a se vedea subsecțiunea 2.1.2), atunci pentru o caracteristică de tip lungime, evaluată

Scanned with OKEN Scanner

la N [pixel], conversia în unități metrice conduce la valoarea $l = N/n$ [inch] = $2.54N/n$ [cm]. În condiții identice, pentru o caracteristică de tip arie evaluată la N [pixel^2], conversia în unități metrice conduce la valoarea $a = N/n$ [inch^2] = $(2.54)^2N/n$ [cm^2]. Odată ce rezultatele măsurătorilor au fost aduse în sistemul metric, se aplică transformările necesare pentru exprimarea valorilor în unități adecvate: mm (mm^2), μm (μm^2), nm (nm^2) etc.

O discuție aparte se impune în cazul imaginilor digitale pentru care nu a fost asigurată aceeași rezoluție pe orizontală și, respectiv, verticală, în sensul că pe axa orizontală rezoluția este n [ppi], iar pe axa verticală – m [ppi]. În această situație, un același număr N de pixeli măsurat pe orizontală și pe verticală conduce la lungimi exprimate în unități metrice diferite, și anume: N/n [inch] = $2.54N/n$ [cm] pe orizontală, și, respectiv, N/m [inch] = $2.54N/m$ [cm] pe verticală. În evaluarea unui perimetru (conform (2.6.3)) apar însă dificultăți, deoarece pentru segmentele plasate oblic formula de calcul prevede corecții care, în mod implicit, se referă la aceeași rezoluție pe cele două axe. Pachetele software performante includ opțiuni pentru a opera cu rezoluții diferite pe cele două axe, formulele de calcul pentru diverși descriptori fiind adaptate automat.

Un alt aspect deosebit de important pentru efectuarea corectă a măsurătorilor în unități metrice îl constituie evaluarea exactă a **puterii de mărire a sistemului optic** prin intermediul căreia se achiziționează imaginea digitală. Mai general spus, pentru a acoperi orice situație creată prin utilizarea a diverse aparaturi de achiziționare a imaginilor digitale, trebuie realizată operația de **calibrare** care stabilește corespondența dintre dimensiunile imaginii microscopice (câmpul ales de pe preparat) și dimensiunile imaginii care face obiectul procesării asistate de calculator.

Exemplul 2.6.2

Ne plasăm în contextul exemplului 2.6.1, considerând aceeași imagine inițială (fig. 2.6.1) și aceeași imagine-mască (fig. 2.6.2). De dată aceasta, măsurătorile sunt efectuate direct în unități metrice, respectiv în micrometri (μm). Se execută operații de măsurare, parametrii urmăriți fiind aria, perimetrul și lungimea structurilor de interes. (Lungimea unei structuri reprezintă, în această situație, latura mare a dreptunghiului de arie minimă în care poate fi înscrisă structura.) Rezultatele măsurătorilor indică existența a 145 de regiuni, corespondența dintre arie și perimetru fiind următoarea:

- două celule cu aria minimă (13.60 μm^2) au perimetrul de 17.11 μm, respectiv 20.83 μm;
- celula cu perimetrul minim (16.07 μm) are o arie de 14.00 μm^2;
- celula cu aria maximă (120.92 μm^2) are un perimetru de 147.36 μm, care este totodată și perimetrul maxim existent în secvența de măsurători.

Valorile în unități metrice obținute în acest exemplu permit evidențierea legăturii cu valorile exprimate în pixeli din exemplul 2.6.1, fiind vorba despre aceleași trăsături. Imaginea din fig. 2.6.1 are 1200 pixeli pe axa orizontală și 852 pixeli pe axa verticală, prezentând aceeași rezoluție, de 300 ppi (= 118 pixeli/cm), pe ambele axe. Totodată, imaginea are dimensiunile reale 10.17 cm x 7.22 cm, fiind achiziționată printr-un sistem optic cu putere de mărire de 400 de ori. Referindu-ne acum la preparatul microscopic, câmpul real corespunzător imaginii din fig. 2.6.1 are dimensiunile 254.25 μm

Scanned with OKEN Scanner

(= 10.17 cm/400) x 180.5 μm (= 7.22 cm/400). Cu referire la acest câmp real, se poate spune că unui micrometru îi corespund 4.719 pixeli pe fiecare din cele două axe. Astfel, de exemplu, pentru cele două celule cu arie minimă, putem converti valorile exprimate în pixeli din exemplul 2.6.1 în unităţi metrice:

- aria de 303 pixeli^2 înseamnă 303 pixeli^2/(4.7192 pixeli^2/μm^2) = 13.60 μm^2;
- perimetrul de 80.77 pixeli înseamnă 80.77 pixeli/(4.719 pixeli/μm) = 17.11 μm;
- perimetrul de 98.33 pixeli înseamnă 98.33 pixeli/(4.719 pixeli/μm) = 20.83 μm.

Valorile obţinute pentru lungimea structurilor celulare sunt cuprinse în gama 5.93 μm – 27.05 μm. Valoarea medie a lungimii, calculată pentru toate cele 145 de structuri măsurate, este de 9.90 μm. Acest rezultat este în deplină concordanţă cu valorile cunoscute din literatură, ţinând cont de faptul că celulele predominante din imagine sunt fibrocite şi fibroblaste (ale căror dimensiuni se încadrează între 10-25 μm), pe lângă acestea existând însă şi alte tipuri de celule (limfocite, mastocite, macrofage).

2.6.3. DESCRIPTORI BIFAZICI SPECIFICI MORFOMETRIEI CELULARE

Parametrii bifazici sunt proprii aprecierilor cantitative în morfologia normală şi patologică, deoarece sunt definiţi prin referire directă la componentele structurale ale unei celule. Dintre aceştia, cei mai utilizaţi sunt:

- $raportul\ nucleo\text{–}citoplasmatic = \dfrac{aria\ nucleului}{aria\ citoplasmei}$;
- $raportul\ nuclear\text{–}nucleolar = \dfrac{aria\ nucleului}{aria\ nucleolului}\left(= \dfrac{aria\ nucleului}{aria\ nucleolilor}\right)$;
- $diferenta\ ariilor = aria\ citoplasmei - aria\ nucleului$;
- $excentricitatea\ nucleului = \dfrac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{D_1}$,

unde x_1, y_1 sunt coordonatele centrului de greutate al citoplasmei, x_2, y_2 sunt coordonatele centrului de greutate al nucleului, iar D_1 este diametrul echivalent al citoplasmei.

2.6.4. UTILIZAREA PARAMETRILOR GEOMETRICI BIDIMENSIONALI ÎN INTERPRETĂRI STEREOLOGICE

Descriptorii geometrici standard ce caracterizează imaginea bidimensională specifică microscopiei optice nu sunt direct aplicabili în investigarea unei structuri, în realitatea sa tridimensională. Totuşi, aceşti descriptori pot fi utilizaţi într-o manieră indirectă pentru estimarea unor caracteristici tridimensionale, atunci când ne plasăm în cadrul ipotezelor şi interpretărilor stereologice. Caracteristicile tridimensionale pot fi astfel estimate prin **stereologie** (*eng.* stereology), conform căreia procentul (din aria întregii entităţi secţionate) pe care îl deţine aria unei anumite structuri este egal cu procentul (din întregul volum al entităţii tridimensionale) pe care îl deţine volumul structurii respective. Astfel, în baza principiilor utilizate de stereologie, măsurătorile în plan bidimensional pot oferi informaţii

Scanned with OKEN Scanner

tridimensionale despre structurile de interes. **Principiile stereologiei** garantează următorul tip de **corespondență** între descriptorii geometrici bidimensionali (2D) și cei tridimensionali (3D), după cum urmează:

- volumul (3D) corespunde unei arii (2D);
- suprafața (3D) corespunde unui contur (2D);
- lungimea (3D) corespunde unui punct ca element cuantificabil (2D).

Observație. În vederea evitării unor eventuale confuzii între termenii suprafață și arie, care în limbajul obișnuit au aceeași semnificație, precizăm că, în contextul stereologiei, termenul de suprafață are în exclusivitate semnificație tridimensională (adică suprafață laterală a unui corp), iar termenul de arie are în exclusivitate semnificație bidimensională (adică aria delimitată de un contur în plan).

Metodele stereologice operează cu estimări asupra proporției (procentului) unei trăsături, prin raportarea strictă la unitatea de volum, adică asupra densităților. În acest sens se utilizează următorii **descriptori**:

- **Densitatea de volum** – V_V (adimensional), definită ca volumul din structura de interes existent în unitatea de volum, sau, echivalent, fracția (procentul) volumului total al structurii de interes dintr-un volum de referință. Pe baza informațiilor bidimensionale se calculează cu formula:

$$V_V = \frac{\text{volumul total al structurii de interes}}{\text{volumul de referinta}} = \frac{\text{aria structurii de interes}}{\text{aria de referinta}}. \qquad (2.6.10)$$

- **Densitatea de suprafață** – S_V [(unitate de arie)/(unitate de volum)], definită ca suprafața structurii de interes în unitatea de volum, sau echivalent, raportul dintre suprafața totală a structurii de interes și volumul de referință. Pe baza informațiilor bidimensionale se calculează cu formula:

$$S_V = \frac{\text{suprafata totala a structurii de interes}}{\text{volumul de referinta}} = \frac{\text{perimetrul structurii de interes}}{\text{aria de referinta}}. \qquad (2.6.11)$$

- **Densitatea de lungime** – L_V [(unitate de lungime)/(unitate de volum)], definită ca lungimea trăsăturii liniare de interes în unitatea de volum sau, echivalent, ca raportul dintre lungimea totală a trăsăturii liniare de interes și volumul de referință. Pe baza informațiilor bidimensionale se calculează cu formula:

$$L_V = \frac{\text{lungimea totala a trasaturii liniare de interes}}{\text{volumul de referinta}} = \frac{\text{numarul punctelor de interes}}{\text{aria de referinta}}. \qquad (2.6.12)$$

Observație. Principiile stereologiei și corespondențele evidențiate anterior sunt reflectate în formulele densităților V_V, S_V, L_V, prin simplificarea termenului aferent lungimii pe a treia coordonată (fără apariția explicită în aceste formule a grosimii structurii de interes).

- **Densitatea numerică** – N_V [(număr structuri)/(unitate de volum)], definită ca numărul de structuri de interes în unitatea de volum, sau echivalent, ca raportul dintre numărul total de structuri de interes și volumul de referință. Pe baza informațiilor bidimensionale se calculează cu formula:

Scanned with OKEN Scanner

$$N_V = \frac{\text{numarul total de structuri de interes}}{\text{volumul de referinta}} = \\ = \frac{\text{numarul total de structuri de interes}}{\text{arie de referinta } x \text{ grosime}} = \\ = \frac{\text{numarul de structuri de interes pe unitatea de arie}}{\text{grosime}} \quad , \qquad (2.6.13)$$

unde "grosime" este distanța între două plane paralele consecutive, apreciată pe baza grosimii medii a celulei pe direcția de secționare, astfel încât celulele numărate într-un plan să nu se mai regăsească în următorul plan. (Atragem atenția asupra faptului că aceste plane sunt practicate virtual, la o distanță impusă de dimensiunile celulelor, care distanță nu trebuie confundată cu standardul de secționare de 4 μm, recomandat în elaborarea preparatelor histologice.)

Utilizarea oricăruia din acești indicatori are drept scop evaluarea cantitativă pentru o trăsătură a structurii de interes corespunzătoare unui volum dat, ceea ce înseamnă produsul dintre valoarea volumului dat și valoarea indicatorului. Cu alte cuvinte, indicatorul joacă rolul numeric al unei densități, fapt ce atrage și denumirea tuturor respectivilor indicatori drept "densitate de".

Rezultatele ce se pot obține prin utilizarea acestor indicatori au relevanță în aprecierea tridimensională a structurii de interes numai în condițiile în care preparatele microscopice analizate au fost realizate prin alegerea unor secțiuni reprezentative, respectând regulile de **eşantionare** impuse de stereologie. Problema eșantionării pentru măsurătorile stereologice (abordată succint în capitolul 1, subsecțiunea 1.3.2) poate fi aprofundată consultând suplimentar o serie de materiale (Howard, Reed, 1998, Russ, Dehoff, 2000, Campbell, 2002 a, e, Slomianka, 2002, Systems Planning and Analysis, Inc., 2002).

Exemplul 2.6.3

Ne plasăm în contextul exemplului 2.6.2, cu măsurătorile asupra imaginii din fig. 2.6.1 efectuate direct în unități metrice, și anume în micrometri (μm). Considerând că preparatul utilizat satisface ipotezele de lucru specifice stereologiei (în sensul că prezintă o oarecare regularitate tridimensională), putem calcula următorii descriptori stereologici. Totuși, acest tip de regularitate nu este în mod riguros suficient pentru a aplica formulele (2.6.10) – (2.6.13), care consideră că deplasarea, de-a lungul celei de-a treia dimensiuni, nu aduce informații suplimentare în raport cu secțiunea plană avută la dispoziție. Regularitatea din cazul pulpei dentare (ca țesut conjunctiv lax) face ca, în deplasarea de-a lungul celei de-a treia dimensiuni, să existe o similaritate de factură statistică, dar fără o repetare identică a organizării geometrice.

Densitatea de volum, conform (2.6.10), este:

$$V_V = \frac{\sum_{k=1}^{145} A_k}{(10.17\text{cm}/400) \times (7.22\text{cm}/400)} = \frac{3643.13\,\mu\text{m}^2}{254.25\,\mu\text{m}\, x\, 180.5\,\mu\text{m}} = \\ = \frac{3643.13\,\mu\text{m}^2}{45892.125\,\mu\text{m}^2} = 0{,}079 = 7.9\%$$

înțelegându-se că 7.9% din orice volum tisular considerat este ocupat de celule.

Scanned with OKEN Scanner

Densitatea de suprafață, conform (2.6.11), este:

$$S_V = \frac{\sum_{k=1}^{145} P_k}{(10.17\text{cm}/400)\times(7.22\text{cm}/400)} = \frac{4449.78\mu\text{m}}{254.25\,\mu\text{m}\,x\,180.5\,\mu\text{m}} =$$

$$= \frac{4.44978\text{mm}}{0.045892125\text{mm}^2} = 97\frac{\text{mm}^2}{\text{mm}^3}$$

înțelegându-se că într-un mm^3 de pulpă dentară există 97 mm^2 înveliș celular.

Densitatea numerică, conform (2.6.13), este:

$$N_V = \frac{\frac{145}{(10.17\text{cm}/400)\times(7.22\text{cm}/400)}}{10\mu\text{m}} = \frac{\frac{145}{0.045892125\text{mm}^2}}{0.01\text{mm}} = 315958\frac{\text{celule}}{\text{mm}^3}$$

înțelegându-se că într-un mm^3 de pulpă dentară există 315958 celule. Valoarea de 10 μm utilizată în calcul reprezintă parametrul "grosime" în sensul formulei (2.6.13) și a fost determinat numeric în exemplul 2.6.2, ca valoare medie a lungimii (9.90 μm ≈ 10 μm).

Scanned with OKEN Scanner

Scanned with OKEN Scanner

CAPITOLUL 3

INTERPRETAREA MATEMATICĂ A REZULTATELOR MĂSURĂTORILOR

Măsurătorile care se referă la trăsăturile cantitative ale mai multor entități de același tip, dintr-un câmp microscopic studiat, furnizează rezultate ce caracterizează atât fiecare entitate în parte, cât și ansamblul acestora. Având în vedere neuniformitatea rezultatelor provenite din măsurarea entităților respective, ansamblul nu poate fi caracterizat decât în **termeni statistici**. Astfel, informații statistice despre un câmp vizează într-o oarecare măsură mulțimea tuturor entităților de același tip (care este cu mult mai vastă decât numărul entităților incluse în câmpul analizat). În condițiile când câmpul analizat este ales după anumite criterii, spre a-i asigura reprezentativitatea în raport cu întreaga mulțime a entităților, *concluziile care se formulează pe baza entităților incluse în câmp se pot extrapola pentru mulțimea tuturor entităților*. Mai mult, în numeroase situații, afirmațiile referitoare la mulțimea tuturor entităților necesită investigații pentru mai multe câmpuri considerate reprezentative, *în conformitate cu legile selecției și eșantionării*. Pe parcursul acestui capitol nu ne vom ocupa de criteriile ce asigură reprezentativitatea unui anumit câmp, și nici de modalitățile de apreciere a validității concluziilor extrapolate. În cadrul acestei lucrări, care se concentrează asupra tehnicilor de morfometrie computerizată, am considerat oportună prezentarea acelor **noțiuni fundamentale** din **teoria probabilităților** și **statistică**, ce sunt necesare pentru o interpretare globală a rezultatelor măsurătorilor la nivel de câmp microscopic.

3.1. NOȚIUNI DESPRE VARIABILELE ALEATOARE

Interpretarea în sens statistic a unei/unor caracteristici cantitative asociată/asociate unui ansamblu de entități se bazează pe conceptele de **probabilitate** și **variabile aleatoare**, motiv pentru care secțiunea curentă oferă un breviar teoretic ce permite operarea riguroasă cu respectivele concepte. Expunerea este construită pe formulările matematice cele mai simple și acordă o atenție deosebită comentării acestor formulări în sens intuitiv. Față de noțiunile de probabilitate și variabilă aleatoare cu **număr finit de valori**, ce sunt în general cunoscute în

urma studierii matematicii de nivel mediu (liceal), au fost introduse suplimentar noțiunile de variabilă aleatoare cu **număr infinit de valori**, de tip **discret** și **continuu**. Prezentarea acestora din urmă a fost abordată ca o generalizare firească, ușor de parcurs, a variabilelor aleatoare cu număr finit de valori. Precizăm că introducerea conceptelor suplimentare este impusă de înțelegerea caracterului continuu al informației investigate, în sensul că *procesul de măsurare poate furniza orice valori dintr-un anumit domeniu, oricât de apropiate unele de altele*. Acest breviar teoretic nu tratează decât situația variabilelor aleatoare monodimensionale, pe care o considerăm suficientă pentru crearea unui cadru matematic în concordanță cu obiectivele de fond ale lucrării. Pe principiile construite, cititorul poate aprofunda studiul și, totodată, poate acumula cunoștințe privind variabilele aleatoare multidimensionale (care acoperă simultan mai multe caracteristici cantitative), consultând referințe bibliografice adecvate secțiunii de față (Mihoc, Micu, 1971, Șabac et al., 1983, Iosifescu et al., 1985).

3.1.1. NOȚIUNEA DE PROBABILITATE

Considerăm o experiență și un eveniment A corespunzător acestei experiențe. Să repetăm această experiență de n ori, în condiții identice, notând cu α numărul de realizări ale evenimentului A și cu n-α numărul de nerealizări ale lui A.

Numărul $\nu_n = \alpha/n$ poartă numele de **frecvență**. Numărul α poate varia de la 0 la n inclusiv. Numărul α ia valoarea 0 atunci când evenimentul A nu s-a realizat niciodată în cadrul celor n repetări consecutive ale experienței. Din contră, dacă în cadrul celor n repetări consecutive ale experienței, evenimentul A s-a realizat în toate experiențele, rezultă $\alpha = n$. În toate celelalte cazuri $0 < \alpha < n$. Constatăm așadar că $0 \leq \nu_n \leq 1$, oricare ar fi numărul de experiențe n care se fac.

Experimental, constatăm că pentru multe fenomene frecvența ν_n, cu n crescând necontenit, se apropie de o valoare constantă. O atare proprietate poartă numele de **lege a numerelor mari**.

Fie A și B două evenimente referitoare la aceeași experiență. Dacă, din motive de perfectă simetrie, putem afirma că ambele evenimente au aceeași șansă de a fi realizate, spunem că evenimentele sunt egal posibile.

Probabilitatea unui eveniment este egală cu raportul dintre numărul cazurilor egal posibile ce realizează evenimentul și numărul tuturor cazurilor egal posibile. Cu alte cuvinte, vom spune că probabilitatea evenimentului A este egală cu raportul dintre numărul m al cazurilor favorabile realizării evenimentului A și numărul n al cazurilor egal posibile. Vom nota aceasta prin $P(A)=\frac{m}{n}$.

3.1.2. VARIABILE ALEATOARE CU UN NUMĂR FINIT DE VALORI

În numeroase experiențe se întâlnesc mărimi care au valori ce se schimbă sub influența unor factori întâmplători. Pentru început, ne concentrăm atenția asupra acelor mărimi care pot lua doar un **număr finit de valori**. Din cauza factorilor întâmplători, fiecare din mărimile considerate poate lua diferite valori în efectuarea repetată a experienței, chiar dacă toate condițiile rămân neschimbate de la o experiență la alta. De aceea, aceste mărimi vor fi numite **variabile aleatoare** (întâmplătoare). Întrucât variabila aleatoare poate lua doar un număr finit de valori, variabila aleatoare va avea implicit și un caracter **discret**. Sensul complet al caracterizării variabilei ca discretă va fi comentat în subsecțiunea 3.1.3.

Scanned with OKEN Scanner

Evident, pentru definirea unei variabile aleatoare, trebuie să cunoaştem valorile pe care aceasta le poate lua. Totuşi, cunoaşterea acestor valori rămâne insuficientă, deoarece fiecare valoare este luată sub influenţa unor factori întâmplători, astfel încât unele valori pot apărea mult mai des decât celelalte. În consecinţă, pentru precizarea completă a variabilei aleatoare, fiecărei valori pe care aceasta o poate lua trebuie să i se asocieze **probabilitatea** cu care este luată valoarea respectivă.

Pentru o variabilă aleatoare f se utilizează notaţia:

$$f\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}, \tag{3.1.1}$$

unde pe primul rând al tabloului sunt trecute, în ordine crescătoare, valorile posibile ale variabilei aleatoare $x_i, i=1,\dots,n$, iar pe rândul al doilea sunt trecute probabilităţile $0 \le p_i \le 1$, $i=1,\dots,n$, cu care pot apărea fiecare din aceste valori. Având în vedere faptul că variabila aleatoare ia cu certitudine una din valorile $x_1,\dots,x_n$ (după cum s-a convenit anterior), probabilităţile aferente vor satisface condiţia $p_1+p_2+\dots+p_n=1$.

Tabloul construit în această manieră defineşte **distribuţia** sau **repartiţia de probabilitate** a variabilei aleatoare f. Acestui tablou i se poate asocia o funcţie

$$F:\mathbf{R}\to[0,1],\ F(x)=P(\{f<x\})=\begin{cases} 0, x\in(-\infty,x_1), \\ p_1, x\in[x_1,x_2), \\ p_1+p_2, x\in[x_2,x_3), \\ \dots\dots\dots \\ p_1+p_2+\dots+p_{n-1}, x\in[x_{n-1},x_n), \\ 1, x\in[x_n,\infty), \end{cases} \tag{3.1.2}$$

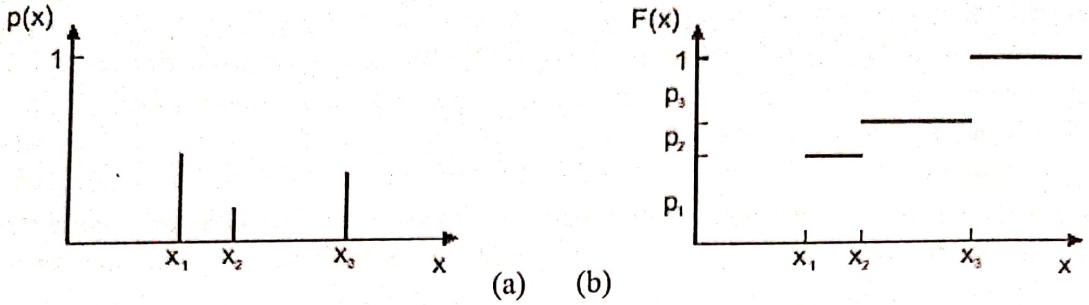
care se numeşte **funcţia de repartiţie** sau **legea de repartiţie** a variabilei aleatoare f. Grafic vorbind, $F(x)$ este o funcţie scară, cu $n+1$ trepte crescătoare, dintre care prima are valoarea 0, iar ultima – valoarea 1.

Exemplul 3.1.1

Pentru ilustrarea celor afirmate anterior, considerăm cazul unei variabile aleatoare care ia trei valori distincte, adică particularizarea $n=3$ în formula funcţiei de repartiţie. În reprezentarea grafică din fig. 3.1.1(b) se observă efectul cumulativ al funcţiei de repartiţie (3.1.2), care este o funcţie scară cu 4 valori: $0, p_1, p_1+p_2, p_1+p_2+p_3=1$ (prezentând salturile p_1, p_2, p_3 corespunzătoare probabilităţilor reprezentate grafic în fig. 3.1.1(a)). Având în vedere definiţia funcţiei F, sunt evidente următoarele trei proprietăţi ale acesteia:

1. F este monoton crescătoare – deoarece, cu cât valoarea x considerată este mai mare, cu atât este mai mare şi valoarea funcţiei.
2. $\lim\limits_{x\to-\infty} F(x)=0$ - deoarece, pentru valori extrem de mici ale lui x, ne plasăm cu certitudine în intervalul $(-\infty,x_1)$.

3. $\lim_{x\to\infty} F(x)=1$ - deoarece, pentru valori extrem de mari ale lui x, ne plasăm cu certitudine în intervalul $[x_n,\infty)$.



(a) (b)

Fig. 3.1.1. Reprezentare grafică asociată unei variabile aleatoare cu trei valori distincte: (a) distribuția de probabilitate; (b) funcția (legea) de repartiție.

Modul de definire a funcției de repartiție F a variabilei aleatoare f permite calculul imediat al probabilității ca f să ia valori într-un anumit interval. De exemplu, probabilitatea ca valoarea luată de variabila aleatoare f să fie mai mare sau egală decât a $(x_j \le a < x_{j+1})$ și mai mică strict decât b $(x_{j+1} < x_k \le b < x_{k+1})$ este

$$P(a \le f < b) = F(b) - F(a) = (p_1 + \ldots + p_k) - (p_1 + \ldots + p_j) = p_{j+1} + \ldots + p_k. \quad (3.1.3)$$

Se numește **medie** (*eng.* mean) a variabilei aleatoare f și se notează $M(f)$ valoarea constantă (numărul) definită prin:

$$M(f) = p_1x_1 + p_2x_2 + \ldots + p_nx_n = \sum_{i=1}^{n} p_ix_i \,. \quad (3.1.4)$$

Acest număr caracterizează variabila aleatoare f cu privire la modul de **grupare** a valorilor pe care le poate lua f.

Se numește **dispersie** sau **varianță** (*eng.* variance) a variabilei aleatoare f și se notează $D(f)$ valoarea constantă (numărul) definită prin:

$$\begin{aligned} D(f) &= M[(f - M(f))^2] = \\ &= p_1(x_1 - M(f))^2 + p_2(x_2 - M(f))^2 + \ldots + p_n(x_n - M(f))^2 = \\ &= \sum_{i=1}^{n} p_i(x_i - M(f))^2. \end{aligned} \quad (3.1.5)$$

Acest număr caracterizează variabila aleatoare f cu privire la modul de **împrăștiere** a valorilor pe care le poate lua f.

3.1.3. VARIABILE ALEATOARE CU UN NUMĂR INFINIT DE VALORI

Conceptul de variabilă aleatoare prezentat anterior este limitat de faptul că aceasta ia un număr finit de valori. Având în vedere situațiile reale, care, în mod frecvent, se caracterizează

Scanned with OKEN Scanner

printr-un **număr infinit de valori** luate de o variabilă aleatoare, conceptul este generalizat corespunzător, după cum se detaliază mai jos.

Se iau în considerare două situații distincte, și anume variabila aleatoare **discretă** și, respectiv, variabila aleatoare **continuă**. În ambele situații, numărul de valori luate fiind infinit, descrierea de tip tablou (3.1.1) nu mai este adecvată, deoarece ar însemna utilizarea unui număr infinit de coloane în tablou. Din acest motiv, variabila aleatoare se va defini direct prin intermediul **funcției** sau **legii de repartiție**. Diferența între cele două situații (discret, respectiv continuu) este dată de "gradul de apropiere" dintre valorile ce le poate lua variabila aleatoare și, implicit, de "multitudinea" acestor valori.

Evitând o formulare matematică dificilă, vom spune că **variabila aleatoare** este **de tip continuu** atunci când valorile pe care le ia pot fi oricât de mult apropiate una de cealaltă, mulțimea valorilor fiind, de fapt, mulțimea numerelor reale **R**. Altfel spus, mulțimea valorilor este densă, în sensul că între oricare două valori există întotdeauna încă o valoare. În cazul unei **variabile aleatoare discrete**, valorile luate de aceasta sunt doar anumite numere din **R**, distanțate (de unde provine și asocierea cu termenul "discret") similar modului de plasare a numerelor naturale **N** în cadrul lui **R**. Definirea ambelor tipuri de variabile aleatoare cu ajutorul funcției sau legii de repartiție se poate realiza într-o manieră comună, care generalizează cele prezentate pentru cazul unui număr finit de valori.

Fie f o **variabilă aleatoare cu număr infinit de valori**, de tip discret sau continuu. Un exemplu tipic pentru cazul continuu îl constituie caracteristica de tip lungime a fibrei musculare striate, care poate lua orice valoare dintr-un interval cuprins între câțiva mm și zeci de cm. Să considerăm x o valoare reală oarecare. Funcția ce desemnează probabilitatea cu care f ia valori strict mai mici decât x

$$F:\mathbf{R}\to[0,1],\ F(x)=P(\{f<x\}) \tag{3.1.6}$$

se numește **funcția de repartiție** sau **legea de repartiție** a variabilei aleatoare f.

Având în vedere semnificația de probabilitate a funcției F, sunt evidente următoarele trei proprietăți ale acesteia:

1. F este monoton crescătoare – deoarece, cu cât valoarea x considerată este mai mare, cu atât este mai mare și probabilitatea ca $f<x$.
2. $\lim_{x\to-\infty} F(x)=0$ - deoarece, pentru o valoare extrem de mică a lui x, cu certitudine nu avem satisfăcută condiția $f<x$, adică probabilitatea ca $f<x$ este 0.
3. $\lim_{x\to\infty} F(x)=1$ - deoarece, pentru o valoare extrem de mare a lui x, cu certitudine avem satisfăcută condiția $f<x$, adică probabilitatea ca $f<x$ este 1.

În cazul (frecvent întâlnit) când, din natura problemei, se știe că variabila aleatoare nu poate lua valori decât într-un interval finit $[v_{\min}, v_{\max}]$, funcția de repartiție F va satisface:

$$F(x)=0, x\in(-\infty, v_{\min});\ F(x)=1, x\in[v_{\max},\infty). \tag{3.1.7}$$

Modul de definire a funcției de repartiție F a variabilei aleatoare f permite calculul imediat al probabilității ca f să ia valori într-un anumit interval. De exemplu, probabilitatea ca valoarea luată de variabila aleatoare f să fie mai mare sau egală decât a și mai mică strict decât b este

$$P(a\le f<b)=F(b)-F(a). \tag{3.1.8}$$

Scanned with OKEN Scanner

Tot din modul de definire a funcției de repartiție F a variabilei aleatoare f reiese că F are un caracter **cumulativ**, care se exprimă în mod diferit pentru f de tip discret și pentru f de tip continuu. Evident, variabila aleatoare cu număr infinit de valori, de tip discret, este – comparativ cu cea de tip continuu - mai apropiată de situația numărului finit de valori, tratată în subsecțiunea 3.1.2. Drept urmare, discutarea caracterului cumulativ al lui F se va face mai întâi pentru variabila de tip discret și, ulterior, pentru cea de tip continuu.

- **CAZUL VARIABILEI ALEATOARE DE TIP DISCRET**

Presupunem că variabila aleatoare f ia valorile $x_1 < \ldots < x_n < \ldots (x_\infty)$, cu probabilitățile $p_1, \ldots, p_n, \ldots, (p_\infty)$, $\sum_{n=1}^{\infty} p_n$, unde s-a folosit notația x_∞, p_∞ pentru a atrage explicit atenția asupra faptului că variabila ia un număr infinit de valori. Precizăm că, din punctul de vedere al notațiilor matematice riguroase, această manieră de notare este abuzivă, ea fiind utilizată însă cu buna intenție de a sublinia un număr infinit de valori.

Informațiile anterioare asupra modului în care ia valori variabila aleatoare discretă pot fi simbolizate într-o manieră rezumativă prin notația $f = \begin{pmatrix} x_n \\ p_n \end{pmatrix}_{n \in I}$, care desemnează **distribuția de probabilitate** a variabilei aleatoare și generalizează notația de tip tablou cu număr finit de valori (3.1.1), pentru un număr infinit de valori, indexate după mulțimea I (care pentru simplitate poate fi privită drept mulțimea numerelor naturale **N**).

Concret, **funcția de repartiție** a variabilei aleatoare f este definită mai jos,

$$F : \mathbf{R} \to [0,1],$$

$$F(x) = P(\{f < x\}) = \begin{cases} 0, x \in (-\infty, x_1), \\ p_1 + p_2 + \ldots + p_{n-1} = \sum_{i=1}^{n-1} p_i, x \in [x_{n-1}, x_n), n = 2,3,\ldots,(\infty), \end{cases} \tag{3.1.9}$$

observându-se cu ușurință generalizarea modului de scriere din cazul variabilei aleatoare cu un număr finit de valori, tratat în subsecțiunea 3.1.2. De asemenea, din punct de vedere grafic, $F(x)$ este tot o funcție scară, dar, spre deosebire de subsecțiunea precedentă, aici există un număr infinit de trepte.

Noțiunile de medie și dispersie, introduse pentru variabilele aleatoare cu număr finit de valori, se generalizează după cum urmează:

Se numește **medie** a variabilei aleatoare f și se notează $M(f)$ valoarea constantă (numărul) definită prin:

$$M(f) = p_1 x_1 + p_2 x_2 + \ldots + p_n x_n + \ldots + (p_\infty x_\infty) = \sum_{n=1}^{\infty} p_n x_n \,. \tag{3.1.10}$$

Acest număr caracterizează variabila aleatoare f cu privire la modul de **grupare** a valorilor pe care le poate lua f.

Se numește **dispersie** sau **varianță** a variabilei aleatoare f și se notează $D(f)$ valoarea constantă (numărul) definită prin:

Scanned with OKEN Scanner

$$D(f)=M[(f-M(f))^2]=$$
$$=p_1(x_1-M(f))^2+p_2(x_2-M(f))^2+\ldots+$$
$$+p_n(x_n-M(f))^2+\ldots+\left(p_\infty(x_\infty-M(f))^2\right)= \qquad (3.1.11)$$
$$=\sum_{n=1}^{\infty}p_n(x_n-M(f))^2.$$

Acest număr caracterizează variabila aleatoare f cu privire la modul de **împrăştiere** a valorilor pe care le poate lua f.

- **CAZUL VARIABILEI ALEATOARE DE TIP CONTINUU**

În această situaţie, datorită modului continuu în care F poate lua valori, caracterul cumulativ al lui F nu mai poate fi exprimat prin sume (ca în cazul discret), ci necesită o exprimare de tip integral. De fapt, datorită densităţii valorilor pe care le poate lua f, nu se mai poate vorbi despre o distribuţie a variabilei aleatoare f cu semnificaţia $f=\begin{pmatrix}x_n\\p_n\end{pmatrix}_{n\in I}$ din cazul discret.

Astfel, se introduce **funcţia** $\rho:\mathbf{R}\to[0,\infty)$ integrabilă pe **R**, care asigură satisfacerea egalităţii:

$$F(x)=\int_{-\infty}^{x}\rho(\xi)d\xi, \qquad (3.1.12)$$

şi care se numeşte **densitate de repartiţie** sau **densitate de probabilitate** a variabilei aleatoare f. Caracterul cumulativ al lui F înseamnă, în termeni grafici, faptul că pentru orice valoare x, $F(x)$ exprimă valoarea ariei cuprinse între axa absciselor şi graficul funcţiei $\rho(\xi)$, situată la stânga abscisei $\xi=x$ (a se vedea semnificaţia geometrică de arie a integralei). Evident, $F(x)$ nu mai poate avea forma unei funcţii scară, ca în cazul discret. Totuşi funcţia $F(x)$ fiind crescătoare, poate fi interpretată ca fiind alcătuită din trepte foarte înguste, de valoare $\rho(x)$ corespunzând unor intervale de lungime infinitesimală dx.

Utilizarea conceptului de densitate de repartiţie se justifică prin aceea că majoritatea legilor (funcţiilor) de repartiţie se definesc analitic şi sunt însoţite de reprezentări grafice relevante, ce exploatează funcţia $\rho(\xi)$ sau, echivalent ca notaţie, $\rho(x)$. Mai mult, această funcţie este utilizată şi în definirea valorilor constante (numere) de tip medie şi dispersie.

Se numeşte **medie** a variabilei aleatoare f şi se notează $M(f)$ valoarea constantă (numărul) definită prin:

$$M(f)=\int_{-\infty}^{+\infty}x\,dF(x)=\int_{-\infty}^{+\infty}x\rho(x)dx. \qquad (3.1.13)$$

Acest număr caracterizează variabila aleatoare f cu privire la modul de **grupare** a valorilor pe care le poate lua f.

Se numeşte **dispersie** sau **varianţă** a variabilei aleatoare f şi se notează $D(f)$ valoarea constantă (numărul) definită prin:

$$D(f)=\int_{-\infty}^{+\infty}(x-M(f))^2\,dF(x)=\int_{-\infty}^{+\infty}(x-M(f))^2\rho(x)dx. \qquad (3.1.14)$$

Scanned with OKEN Scanner

Acest număr caracterizează variabila aleatoare f cu privire la modul de **împrăştiere** a valorilor pe care le poate lua f.

Observaţie. Calculul mediei şi dispersiei pun şi ele în evidenţă trecerea de la exprimarea cumulativă de tip sumă – specifică pentru cazul discret – la exprimarea cumulativă de tip integral – specifică pentru cazul continuu.

În contextul lucrării noastre ne vom limita la prezentarea a două **legi de repartiţie de tip continuu**, întâlnite frecvent în studiile de natură statistică, şi anume **legea uniformă** şi **legea normală**.

Exemplul 3.1.2

Se spune că o variabilă aleatoare f are o **funcţie de repartiţie uniformă** de parametri α,β, cu $\alpha<\beta$ şi se notează prescurtat $f\sim U(\alpha,\beta)$, dacă

$$\rho(x)=\begin{cases}\dfrac{1}{\beta-\alpha}, & x\in[\alpha,\beta],\\ 0, & x\in R\setminus[\alpha,\beta].\end{cases} \tag{3.1.15}$$

Fig. 3.1.2 arată reprezentările grafice asociate la trei variabile aleatoare uniforme cu parametrii (i) $\alpha=1,\beta=3$ - reprezentare prin linie continuă, (ii) $\alpha=1.5,\beta=2.5$ - reprezentare prin linie întreruptă, (iii) $\alpha=4.5,\beta=5.5$ - reprezentare prin linie-punct. Partea de sus a imaginii prezintă graficele funcţiei densitate de repartiţie $\rho(x)$, iar partea de jos, graficele funcţiei (legii) de repartiţie $F(x)$.

Variabila aleatoare f ia valori numai în intervalul $[\alpha,\beta]$ şi orice valoare din acest interval poate fi luată într-o manieră echiprobabilă, datorită faptului că funcţia $\rho(x)$ este constantă. Cu cât intervalul $[\alpha,\beta]$ este mai mic, cu atât valoarea constantă a funcţiei $\rho(x)$ este mai mare, de aşa manieră încât aria cuprinsă între axa absciselor şi graficul funcţiei $\rho(x)$ este unitară, adică $\int_{-\infty}^{+\infty}\rho(x)dx=1$.

În baza definiţiilor de mai sus, rezultă că **media** şi **dispersia** acestei variabile aleatoare au valorile $M(f)=\dfrac{\alpha+\beta}{2}$ şi, respectiv, $D(f)=\dfrac{(\beta-\alpha)^2}{12}$ cu următoarele semnificaţii concrete:

Media este chiar centrul intervalului $[\alpha,\beta]$ în care variabila aleatoare f poate lua valori. Astfel, mai multe variabile aleatoare uniforme, definite pe intervale diferite, au aceeaşi medie dacă şi numai dacă centrele intervalelor respective coincid (indiferent de extremităţile lor).

Dispersia este dictată de mărimea intervalului $[\alpha,\beta]$ în care variabila aleatoare f poate lua valori. Astfel, mai multe variabile aleatoare uniforme, definite pe intervale diferite, au aceeaşi dispersie dacă şi numai dacă lungimile intervalelor respective coincid (indiferent de extremităţile lor).

Scanned with OKEN Scanner

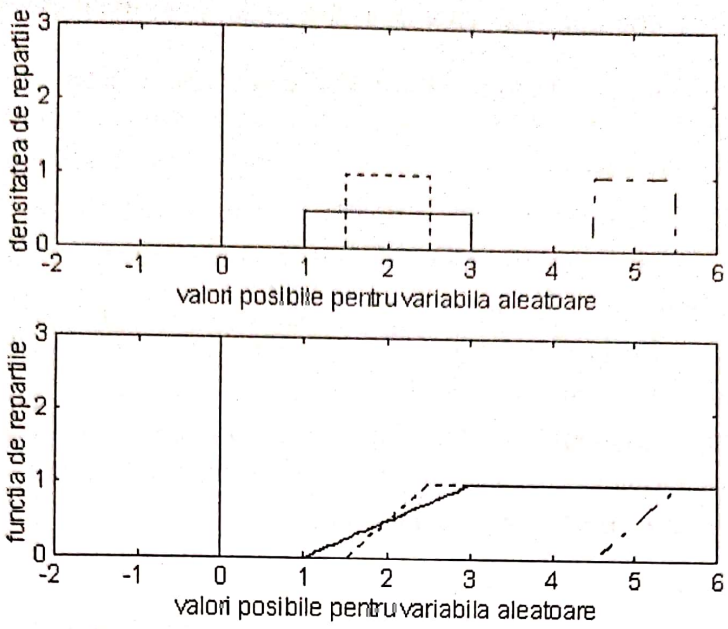


Fig. 3.1.2. Reprezentări grafice asociate la trei variabile aleatoare uniforme cu parametrii
(i) $\alpha=1,\beta=3$ - reprezentare prin linie continuă,
(ii) $\alpha=1.5,\beta=2.5$ - reprezentare prin linie întreruptă,
(iii) $\alpha=4.5,\beta=5.5$ - reprezentare prin linie-punct.

Exemplul 3.1.3

Se spune că o variabilă aleatoare f are o **funcție de repartiție normală** (sau **de tip Gauss**) de parametri m,σ^2 şi se notează prescurtat $f\sim N(m,\sigma^2)$, dacă

$$\rho(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{(x-m)^2}{2\sigma^2}}. \qquad (3.1.16)$$

Fig. 3.1.3 arată reprezentările grafice asociate la trei variabile aleatoare uniforme cu parametrii (i) $m=2,\sigma^2=0.25$ - reprezentare prin linie continuă, (ii) $m=2,\sigma^2=1$ - reprezentare prin linie întreruptă, (iii) $m=4,\sigma^2=0.25$ - reprezentare prin linie-punct. Partea de sus a imaginii prezintă graficele funcției densitate de repartiție $\rho(x)$, iar partea de jos, graficele funcției (legii) de repartiție $F(x)$.

Graficul funcției $\rho(x)$ este simetric în raport cu verticala ce trece prin punctul $x=m$, în care $\rho(x)$ își atinge maximul $\rho(m)=\frac{1}{\sigma\sqrt{2\pi}}$. Cu cât parametrul m are o valoare mai mică, cu atât maximul lui $\rho(x)$ este situat mai la stânga pe grafic. Cu cât parametrul σ^2 are o valoare mai mică, cu atât maximul lui $\rho(x)$ este mai mare, iar creşterea (la stânga lui m), respectiv descreşterea (la dreapta lui m) funcției $\rho(x)$ este mai pronunțată.

Scanned with OKEN Scanner

Această constatare este o consecință imediată a faptului că, indiferent de valorile parametrilor m, σ^2, aria cuprinsă între axa absciselor și graficul funcției $\rho(x)$ este unitară, adică $\int_{-\infty}^{+\infty} \rho(x)dx = 1$. Datorită alurii graficului său, funcția $\rho(x)$ este uneori denumită "clopotul lui Gauss".

În baza definițiilor de mai sus, rezultă că **media** și **dispersia** acestei variabile aleatoare au valorile $M(f) = m$ și, respectiv, $D(f) = \sigma^2$ cu următoarele semnificații concrete:

Media este chiar parametrul m din definiția variabilei aleatoare f. Astfel, mai multe variabile aleatoare normale au aceeași medie dacă și numai dacă au același parametru m (legile de repartiție diferă prin parametrul σ^2).

Dispersia este chiar parametrul σ^2 din definiția variabilei aleatoare f. Astfel, mai multe variabile aleatoare normale au aceeași dispersie dacă și numai dacă au același parametru σ^2 (legile de repartiție diferă prin parametrul m).

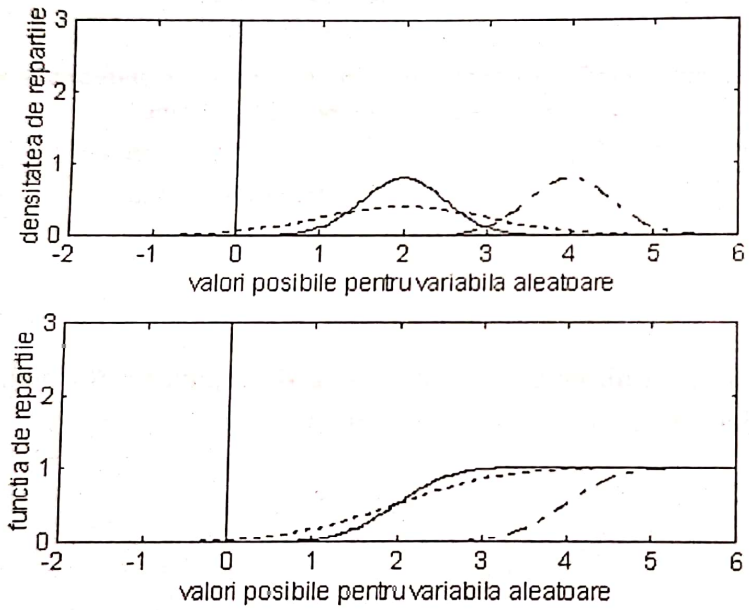


Fig. 3.1.3. Reprezentări grafice asociate la trei variabile aleatoare uniforme cu parametrii

(i) $m = 2, \sigma^2 = 0.25$ - reprezentare prin linie continuă,

(ii) $m = 2, \sigma^2 = 1$ - reprezentare prin linie întreruptă,

(iii) $m = 4, \sigma^2 = 0.25$ - reprezentare prin linie-punct.

3.1.4. CORELAȚIA A DOUĂ VARIABILE ALEATOARE

Se consideră două variabile aleatoare f și g de aceeași natură (discretă sau continuă). Dacă, din toate valorile pe care le ia f și, respectiv, g, se scad mediile corespunzătoare celor două variabile, se obțin tot două variabile aleatoare $f - M(f)$ și, respectiv, $g - M(g)$.

Scanned with OKEN Scanner

Corelația sau **covarianța** (*eng.* covariance) variabilelor aleatoare f și g este un număr notat Cov[f,g] care se definește astfel:

$$\mathrm{Cov}[f,g]=M\left[(f-M(f))(g-M(g))\right], \tag{3.1.17}$$

adică media variabilei aleatoare ce reprezintă produsul dintre variabilele aleatoare $f-M(f)$ și $g-M(g)$. Totodată, numărul definit

$$C_{fg}=\frac{\mathrm{Cov}[f,g]}{\sqrt{D(f)D(g)}} \tag{3.1.18}$$

se numește **coeficient de corelație** sau **factor de corelație** (*eng.* correlation factor) al variabilelor aleatoare f și g. (Notația D desemnează dispersia, respectând convențiile de notație introduse în subsecțiunile 3.1.2, 3.1.3.)

În cazul când $C_{fg}=0$, se spune că variabilele aleatoare f și g sunt necorelate. Indiferent de valorile pe care le iau cele două variabile aleatoare, coeficientul de corelație este un număr cuprins între -1 și 1, adică $-1\le C_{fg}\le 1$. Dacă $C_{fg}=1$ sau $C_{fg}=-1$, între valorile celor două variabile aleatoare f și g există o relație liniară de forma $Af+Bg+C=0$.

În general vorbind, **coeficientul de corelație** (raportul de corelație) reprezintă o măsură a gradului de legătură între cele două variabile aleatoare considerate. Valoarea absolută (modulul) a acestui coeficient apropiată de unitate indică o puternică dependență liniară între cele două variabile aleatoare, f și g. O valoare absolută mică nu înseamnă, în mod obligatoriu, lipsa unei legături între f și g – această legătură poate să existe, dar nu este de formă liniară. O valoare pozitivă a coeficientului de corelație arată că ambele variabile aleatoare f și g variază în același sens, iar o valoare negativă a acestuia arată că variabile aleatoare f și g variază în sens contrar.

După cum se va vedea în secțiunea 3.2, *corelația a două variabile aleatoare joacă un rol important în interpretarea statistică a secvențelor de date obținute prin măsurători.*

3.2. ELEMENTE DE STATISTICĂ

Informațiile furnizate în secțiunea curentă au drept scop dezvoltarea unui punct de vedere **statistic**, fundamentat teoretic, pentru interpretarea în manieră globală a rezultatelor măsurătorilor efectuate asupra unui ansamblu de entități (de obicei situate în același câmp de observație microscopică). În acest scop, am creat premisele diferențierii noțiunilor de **populație statistică** (ca mulțime atotcuprinzătoare a entităților de un anumit tip) și **eșantion** (ca ansamblu finit de entități) și am introdus și caracterizat noțiunea de **variabilă aleatoare empirică** (legată de experimentul de măsurare pe un eșantion). Pentru corecta înțelegere a caracteristicilor unei distribuții empirice, am considerat utilă realizarea comparației cu distribuțiile teoretice de tip Gauss și de tip lege uniformă, care au fost discutate detaliat pe parcursul secțiunii 3.1. Comparația avută în vedere *se limitează la aspecte de natură intuitivă, și nu vizează teoria selecției și estimației statistice* (în sensul validității, la nivelul întregii populații, a rezultatelor obținute pe eșantion).

Tot în contextul analizei statistice a unui eșantion (adică a ansamblului de entități delimitat de câmpul de observație microscopică) se abordează problematica **dependenței stohastice** a două variabile aleatoare empirice, prezentând, ca instrumente specifice de investigație, **metoda corelației** și **metoda regresiei** (Căruntu, 2003 b). Pentru aprofundarea

Scanned with OKEN Scanner

informațiilor care fac obiectul acestei secțiuni, recomandăm următoarea bibliografie (Mihoc, Micu, 1971, Rumșiski, 1974, Șabac et al., 1983, Iosifescu et al., 1985, Văleanu, Hâncu, 1990).

3.2.1. POPULAȚIE STATISTICĂ, EȘANTION SAU SELECȚIE

În limbaj matematic, orice mulțime care face obiectul unei analize statistice se numește **populație statistică**. Elementele constitutive ale unei populații statistice se numesc **unități statistice** sau **indivizi**. Numărul tuturor indivizilor dintr-o populație statistică se numește efectivul total al acelei populații sau volumul populației. Volumul unei populații statistice poate fi finit sau infinit. Analiza statistică poate avea în vedere una sau mai multe **caracteristici** (trăsături) comune tuturor indivizilor ce alcătuiesc populația statistică. O caracteristică se numește **cantitativă** dacă se poate măsura. În caz contrar, caracteristica se numește **calitativă**.

Informațiile privind valorile unei caracteristici nu se culeg de la întreaga populație, ci se consideră la întâmplare o submulțime finită a populației. Această submulțime și, implicit, valorile corespunzătoare caracteristicii studiate poartă denumirea de **eșantion** sau **selecție**. Numărul de elemente conținut de o selecție se numește volumul selecției. Procedeul de extragere a unui eșantion dintr-o populație statistică se numește **sondaj**.

Analiza statistică în raport cu **o singură caracteristică cantitativă** se bazează pe gruparea datelor sub forma unor perechi de valori, primul număr din pereche fiind valoarea caracteristicii, iar cel de al doilea – numărul de indivizi ce prezintă acea valoare a caracteristicii. Dată fiind o populație statistică și o anumită caracteristică cantitativă, totalitatea perechilor create în modul de mai sus constituie o **serie statistică**. De obicei, colectarea informațiilor referitoare la o serie statistică se poate realiza sub forma unui tabel cu două coloane, corespunzând celor două componente ale perechii, adică în prima coloană se introduce valoarea caracteristicii, iar în cea de a doua coloană se introduce numărul de indivizi. În cazul seriilor statistice de dimensiuni mari, deseori se preferă gruparea datelor pe **clase**, în raport cu valorile caracteristicii studiate. Clasele în care se organizează valorile caracteristicii cantitative sunt, din punct de vedere matematic, intervale închise la stânga și deschise la dreapta, de forma $a \le x < b$, a căror lungime se definește de obicei egală.

Noțiunile introduse anterior, pentru o singură caracteristică cantitativă, se pot generaliza adecvat pentru situația a două sau mai multe caracteristici cantitative.

3.2.2. SERII STATISTICE ȘI VARIABILE ALEATOARE EMPIRICE

Să notăm prin φ caracteristica de tip cantitativ pe care vrem să o studiem pentru o selecție de volum N. Din punct de vedere matematic, φ este o **variabilă aleatoare empirică**, care poate lua cel mult n valori diferite (întrucât informația poate fi obținută numai de la cei N indivizi ai selecției concrete la care ne referim). Din acest motiv, o variabilă aleatoare empirică este de tip **discret**. Pentru a acoperi generalitatea, să considerăm că variabila aleatoare empirică φ ia $n \le N$ valori distincte, așezate în ordine crescătoare:

$$x_1 < x_2 < \ldots < x_n, \; n \le N. \tag{3.2.1}$$

Această succesiune de valori distincte provine prin eliminarea termenilor care se repetă din următoarea succesiune ce conține caracteristicile tuturor celor N indivizi:

$$x_1 \le x_2 \le \ldots \le x_N. \tag{3.2.2}$$

Scanned with OKEN Scanner

Cele două succesiuni de valori considerate sunt referite cu numele de serii statistice.

Fiecărei valori x_i, $i=1,\ldots,n$, îi corespunde în mod unic numărul de indivizi din selecție w_i, $i=1,\ldots,n$, pentru care a fost determinată valoarea respectivă. Numărul natural w_i, $i=1,\ldots,n$, se numește **frecvență absolută** de apariție a valorii x_i, $i=1,\ldots,n$, fiind evidentă satisfacerea egalității:

$$\sum_{i=1}^{n} w_i = N. \tag{3.2.3}$$

Numărul rațional $\upsilon_i = \frac{w_i}{N}$, $i=1,\ldots,n$, se numește **frecvență relativă** de apariție a valorii x_i, $i=1,\ldots,n$, fiind evidentă satisfacerea egalității:

$$\sum_{i=1}^{n} \upsilon_i = 1. \tag{3.2.4}$$

Folosind notațiile matematice consacrate, introduse în secțiunea anterioară, variabilei aleatoare empirice φ i se asociază una din următoarele scrieri echivalente care definesc o distribuție empirică:

$$\begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \upsilon_1 & \upsilon_2 & \ldots & \upsilon_n \end{pmatrix} \text{ sau } \begin{pmatrix} x_i \\ \upsilon_i \end{pmatrix}_{i=1,\ldots,n} \text{ sau } \begin{pmatrix} x_i \\ \upsilon_i \end{pmatrix}_{1\le i\le n}. \tag{3.2.5}$$

Funcția de repartiție empirică asociată variabilei aleatoare empirice φ se definește sub forma cunoscută din cazul variabilelor aleatoare cu număr finit de valori:

$$\Phi : \mathbf{R} \to [0,1],\ \Phi(x) = \begin{cases} 0,\ x < x_1, \\ \upsilon_1, x_1 \le x < x_2, \\ \upsilon_1 + \upsilon_2, x_2 \le x < x_3, \\ \ldots \\ \upsilon_1 + \upsilon_2 + \ldots + \upsilon_{n-1}, x_{n-1} \le x < x_n, \\ 1, x_n \le x. \end{cases} \tag{3.2.6}$$

Această funcție are semnificația unei frecvențe relative cumulate. Așa cum s-a discutat în subsecțiunea 3.1.2, graficul funcției $\Phi(x)$ este de tip funcție scară, cu $n+1$ valori în ordine crescătoare (datorită efectului cumulativ pe care îl are $\Phi(x)$).

În cazul când selecția are un volum N mare, pentru a ușura manipularea datelor este utilă gruparea indivizilor pe **clase** sau **intervale**. Astfel, dacă $I=[a,b]$ este un interval ce conține toate cele N valori pe care le poate lua variabila aleatoare φ, atunci acest interval poate fi divizat în $r<n$ intervale (clase):

$$I_1 = [a, a_1),\ I_k = [a_{k-1}, a_k),\ k = 2,\ldots,r-1,\ I_r = [a_{r-1}, b]. \tag{3.2.7}$$

Frecvent, se preferă ca aceste intervale să fie alese egale. Pentru fiecare clasă I_k, $k=1,\ldots,r$, se desemnează o valoare (din acea clasă) $y_k \in I_k$, $k=1,\ldots,r$, ca **reprezentant** al clasei respective (de exemplu, mijlocul intervalului I_k). Astfel, variabila aleatoare empirică φ va lua următoarele r valori distincte așezate în ordine crescătoare:

$$y_1 < y_2 < \ldots < y_r,\ r < n. \tag{3.2.8}$$

Scanned with OKEN Scanner

Fiecărei valori y_k, $k=1,\ldots,r$, îi corespunde în mod unic numărul de indivizi din selecție q_k, $k=1,\ldots,r$, pentru care au fost determinate valori cuprinse în intervalul I_k, $k=1,\ldots,r$. Numărul natural q_k, $k=1,\ldots,r$, se numește **frecvență absolută** de apariție a valorilor din intervalul I_k, $k=1,\ldots,r$, fiind evidentă satisfacerea egalității:

$$\sum_{k=1}^{r} q_k = N. \tag{3.2.9}$$

Numărul rațional $\mu_k = \frac{q_k}{N}$, $k=1,\ldots,r$, se numește **frecvență relativă** de apariție a valorilor din intervalul I_k, $k=1,\ldots,r$, fiind evidentă satisfacerea egalității:

$$\sum_{k=1}^{r} \mu_k = 1. \tag{3.2.10}$$

Folosind notația consacrată cu tabloul (3.1.1), variabilei aleatoare empirice φ i se asociază una din următoarele scrieri echivalente care definesc o distribuție empirică:

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_r \\ \mu_1 & \mu_2 & \cdots & \mu_r \end{pmatrix} \text{ sau } \begin{pmatrix} y_k \\ \mu_k \end{pmatrix}_{k=1,\ldots,r} \text{ sau } \begin{pmatrix} y_k \\ \mu_k \end{pmatrix}_{1\le k\le r}. \tag{3.2.11}$$

Odată ce a fost definită o astfel de variabilă aleatoare, cu ea se va opera fără a mai ține cont de semnificația celor r valori distincte, construite prin procedeul de divizare în clase. În consecință, funcția de repartiție se va defini raportându-ne la valorile $y_1,\ldots,y_r$ și, respectiv, la frecvențele relative $\mu_1,\ldots,\mu_r$, având o formă similară cu cea scrisă mai sus (3.2.11) pentru variabila cu n valori distincte $\begin{pmatrix} x_i \\ \upsilon_i \end{pmatrix}_{1\le i\le n}$.

3.2.3. REPREZENTAREA GRAFICĂ A SERIILOR STATISTICE

Reprezentarea grafică a unei serii este un instrument foarte sugestiv de analiză, deoarece poate oferi o primă interpretare, de factură intuitivă, a datelor. În cele ce urmează, ne vom referi la serii statistice cu **o singură caracteristică cantitativă**. Acestea se reprezintă grafic în raport cu un sistem de axe rectangulare. Alegerea unității pe fiecare dintre axe se realizează în conformitate cu obiectivul propus.

- **REPREZENTAREA CU BATOANE (SEGMENTE VERTICALE)**

Se utilizează pentru o descriere grafică a seriei statistice, cu referire la fiecare valoare individuală pe care o ia variabila aleatoare empirică (spre deosebire de reprezentările grafice descrise mai jos, care grupează valorile pe clase). Această situație este frecvent întâlnită atunci când variabila aleatoare empirică ia un număr mic de valori. Pe axa absciselor sunt trecute punctele reprezentând valorile variabilei; din fiecare punct de acest fel se ridică segmente verticale de lungime egală cu frecvența absolută, sau cu frecvența relativă a valorii respective. Segmentele ridicate se măsoară cu unitatea corespunzătoare axei ordonatelor.

- **HISTOGRAMA**

Se utilizează pentru o descriere grafică a seriei statistice, cu referire la clase de valori în care sunt grupate valorile variabilei aleatoare empirice. De obicei, clasele se aleg ca intervale egale,

Scanned with OKEN Scanner

ceea ce revine la a plasa segmente egale pe axa absciselor (fiecare segment corespunzând unei clase). În cazul claselor egale, lungimea segmentului generic poartă denumirea de amplitudine a clasei. De pe fiecare segment considerat ca bază se ridică un dreptunghi cu înălțimea proporțională cu frecvența absolută, sau cu frecvența relativă a clasei respective. Organizarea pe clase are în vedere numai realizarea desenului, şi nu vizează raportarea analizei statistice la o nouă variabilă aleatoare empirică, definibilă cu ajutorul claselor (în sensul relației (3.2.11)).

- **POLIGONUL FRECVENȚELOR**

Se utilizează pentru seriile în care valorile variabilei aleatoare empirice sunt grupate pe clase. Considerăm o împărțire în segmente egale plasată pe axa absciselor (ca în cazul histogramei). Din mijlocul fiecărui segment (interval) se ridică un segment proporțional cu frecvența (absolută sau relativă) corespunzătoare clasei în cauză. Prin unirea extremităților superioare ale segmentelor verticale se obține poligonul frecvențelor. Organizarea pe clase are în vedere numai realizarea desenului şi nu vizează raportarea analizei statistice la o nouă variabilă aleatoare empirică, definibilă cu ajutorul claselor (în sensul relației (3.2.11)).

Reprezentările grafice de tip histogramă şi poligonul frecvențelor plasează pe axa ordonatelor valori ale frecvențelor absolute sau relative. Într-o manieră similară, se pot construi reprezentări grafice care să plaseze pe ordonată valori cumulative pentru frecvențele absolute sau relative. În această situație se obține un grafic crescător, denumit **histogramă cumulativă** şi, respectiv, **poligonul cumulativ al frecvențelor**. În cazul cumulării frecvențelor relative, acestea pot fi privite drept aproximări ale funcției de repartiție empirică (3.2.6).

Reprezentările grafice pot fi extinse şi în cazul a **două caracteristici cantitative**, atunci când avem de-a face cu variabile aleatoare bidimensionale. În plan orizontal, se vor plasa valorile variabilelor aleatoare (ca perechi de două numere desemnând cele două caracteristici). Pe axa verticală (de obicei denumită cotă) se plasează frecvența absolută sau relativă corespunzătoare fiecărei perechi de valori luată de variabila aleatoare empirică (adică asociată fiecărui punct de coordonate cunoscute din plan). De obicei, planul orizontal se organizează sub forma unei grile generate de segmente egale luate pe cele două axe. Astfel, histograma se va generaliza la o reprezentare tridimensională, locul dreptunghiurilor fiind luat de paralelipipede dreptunghice plasate în celulele grilei. De asemenea, poligonul frecvențelor se va generaliza la o suprafață alcătuită din porțiuni plane, ale cărei secțiuni sunt linii poligonale (linii frânte).

Exemplul 3.2.1

Se consideră imaginea din fig. 2.4.6 reprezentând o etapă de formare a mugurelui dentar (stadiu de clopot), cu dentina în curs de formare. Elementele propuse pentru măsurare sunt localizate la nivelul organului smalțului, fiind vorba, în fapt, de dimensiunile unor celule din epiteliul reticular al organului smalțului. Au fost alese şi marcate 15 celule de tip reticular, selectate din imaginea originală în mod interactiv; aceste celule pot fi observate sub forma regiunilor albe din imaginea binară rezultată (fig. 3.2.1). Prin măsurarea automată a ariilor celor 15 regiuni a rezultat seria statistică:

$$114<190<234<241<276<319<334<337<368<372<396<417<459<695<815.$$

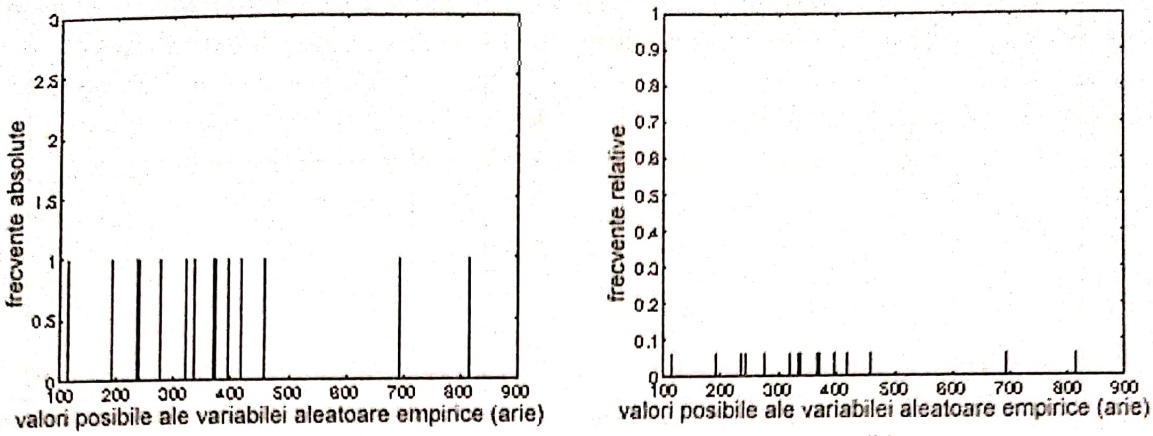
Acesteia îi poate fi ataşată reprezentarea grafică cu batoane (segmente) din fig. 3.2.2. Fig. 3.2.2(a) se referă la frecvențele absolute, unde fiecare valoare a variabilei aleatoare empirice este luată individual. Deoarece toate cele 15 valori sunt distincte, fiecare baton are înălțimea 1. Fig. 3.2.2(b) se referă la frecvențele relative, unde fiecare valoare

Scanned with OKEN Scanner

a variabilei aleatoare empirice este luată individual. Deoarece toate cele 15 valori sunt distincte, fiecare baton are înălțimea 1/15 = 0.0666, adică o frecvență 6,66%.
Considerând o grupare pe clase egale, de amplitudine 100 pixel^2, pentru valorile luate de variabila empirică, în fig. 3.2.3(a) este prezentată histograma aferentă frecvențelor absolute. Din histogramă se observă imediat că rezultatele măsurătorilor, adică valorile luate de variabila aleatoare, sunt dispuse astfel: nici o valoare în clasele [0, 100), [500, 600), [700, 800) şi [900, ∞), o singură valoare în clasele [600, 700) şi [800, 900), două valori în clasele [100, 200) şi [400, 500), trei valori în clasa [200, 300), şase valori în clasa [300, 400). Evident, suma tuturor frecvențelor absolute este 15. Histograma pentru frecvențe relative este prezentată în fig. 3.2.3(b), care diferă de fig. 3.2.3(a) doar prin faptul că pe ordonată sunt marcate valori subunitare, cu semnificație de frecvență relativă. Evident, suma tuturor frecvențelor relative este 1.
Dacă amplitudinea claselor este de 50 în loc de 100 (cum a fost considerat mai înainte), histogramele pentru frecvențele absolute şi relative arată ca în fig. 3.2.4. Astfel, au fost create 16 clase şi modul de distribuire a valorilor variabilei aleatoare empirice pe aceste clase este, evident, diferit comparativ cu situația discutată anterior pentru 8 clase.
Revenim la organizarea pe 8 clase de amplitudine 100. Folosind pe axa absciselor mijloacele intervalelor egale ce desemnează clasele (adică 150, 250, 350, 450, 550, 650, 750, 850) şi pe axa ordonatelor frecvențele absolute corespunzătoare acestor clase (adică 2, 3, 6, 2, 0, 1, 0, 1), se plasează punctele care, unite prin segmente de dreaptă, generează poligonul frecvențelor absolute (fig. 3.2.5(a)). Similar, luând pe axa ordonatelor frecvențele relative, se poate construi poligonul frecvențelor relative, care are aceeaşi alură (fig. 3.2.5(b)).
În fig. 3.2.6 se prezintă histograma cumulativă corespunzătoare frecvențelor relative, care este o funcție crescătoare de tip scară, cu 8 trepte, corespunzătoare celor 8 clase. Remarcăm faptul că, pentru clasele caracterizate prin valoare relativă 0, histograma cumulativă rămâne constantă (la nivelul clasei anterioare). Deşi această reprezentare grafică are un caracter cumulativ, acesta se manifestă la nivelul claselor, şi graficul din fig. 3.2.6 nu trebuie confundat cu graficul legii de repartiție empirică din fig. 3.2.7, trasat pe baza definiției din relația (3.2.6). Acesta din urmă caracterizează variabila aleatoare empirică fără a face apel la gruparea pe clase, constând dintr-o funcție scară, cu 15 trepte (corespunzând numărului total de măsurători – valori distincte în seria statistică). În fig. 3.2.7, de la oricare treaptă la următoarea saltul este constant, datorită faptului că toate valorile variabilei aleatoare empirice au aceeaşi frecvență relativă (de 1/15 = 6,66%, după cum reiese şi din fig. 3.2.2(b)).

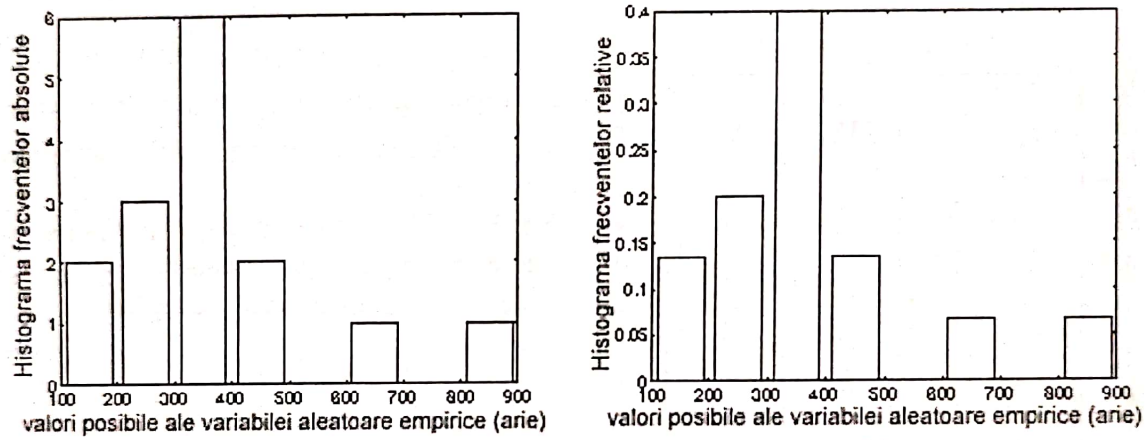
Fig.3.2.1. Imagine-mască pentru măsurătorile statistice.

Scanned with OKEN Scanner



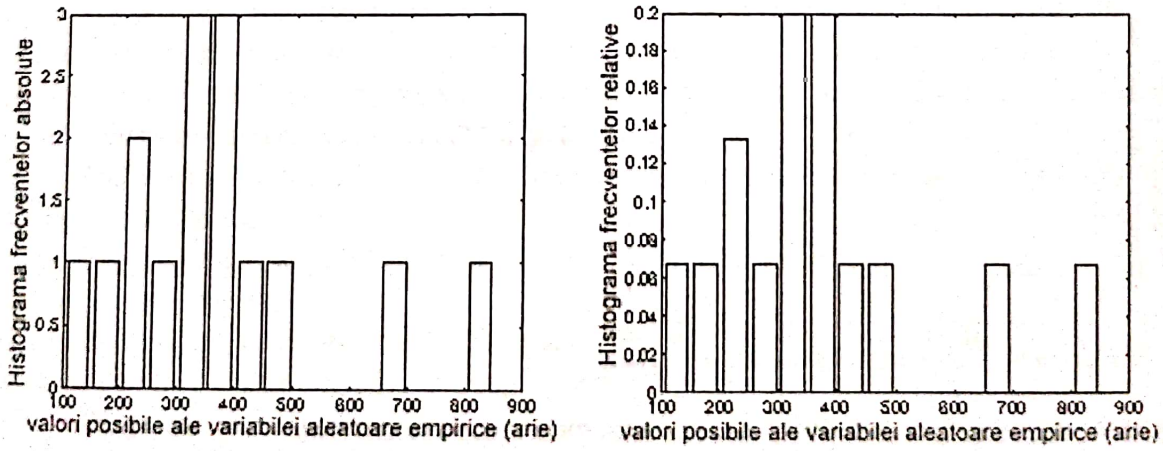
(a) (b)

Fig. 3.2.2. Reprezentare cu batoane (segmente):
(a) pentru frecvențe absolute; (b) pentru frecvențe relative.



(a) (b)

Fig. 3.2.3. Histograma corespunzătoare organizării pe clase de amplitudine 100:
(a) pentru frecvențe absolute; (b) pentru frecvențe relative.



(a) (b)

Fig. 3.2.4. Histograma corespunzătoare organizării pe clase de amplitudine 50:
(a) pentru frecvențe absolute; (b) pentru frecvențe relative.

Scanned with OKEN Scanner

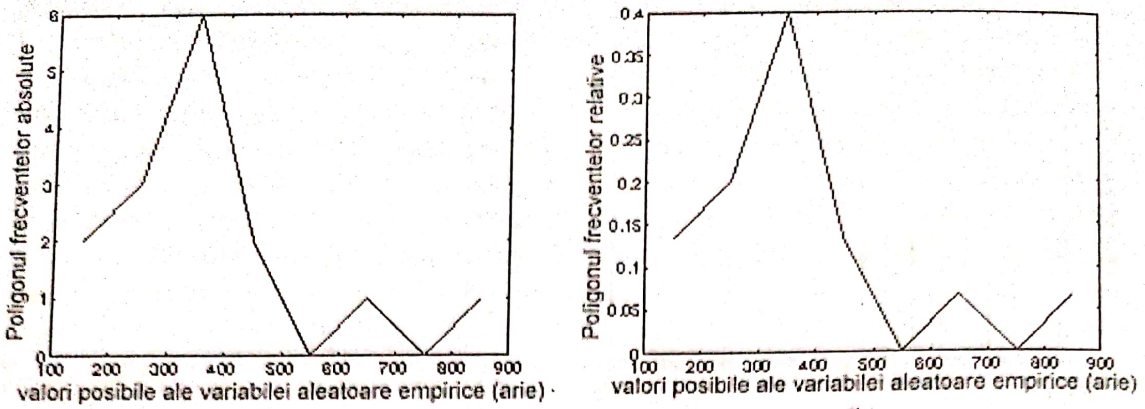


Fig. 3.2.5. Poligonul frecvențelor corespunzătoare organizării pe clase de amplitudine 100: (a) pentru frecvențe absolute; (b) pentru frecvențe relative.

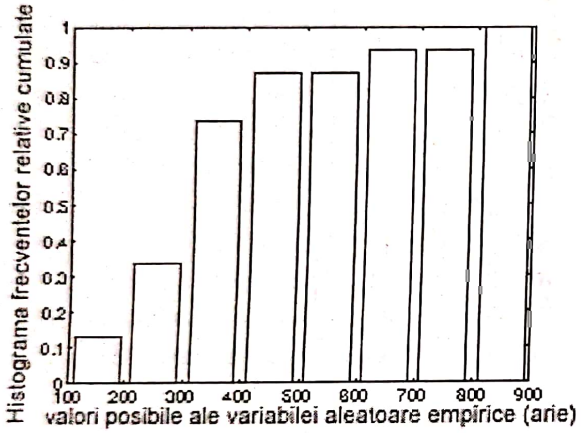


Fig. 3.2.6. Histograma cumulativă pentru frecvențe relative, corespunzătoare organizării pe clase de amplitudine 100.

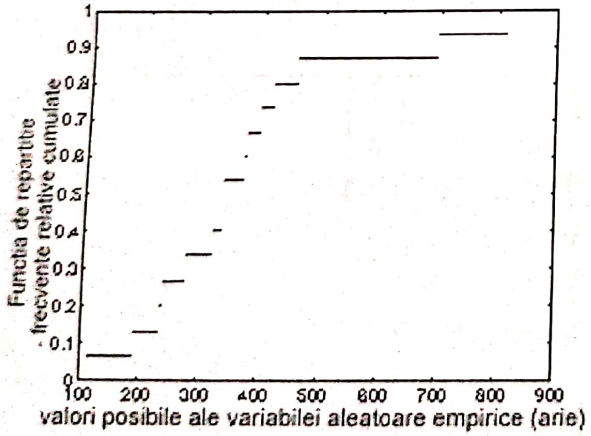


Fig. 3.2.7. Funcția de repartiție empirică.

3.2.4. ELEMENTE CARACTERISTICE ALE SERIILOR STATISTICE

Ca urmare a variabilității mărimilor obținute din măsurători, există o tendință a datelor de a se grupa în jurul unei frecvențe maxime şi o tendință contrară, de împrăştiere față de aceasta.

Scanned with OKEN Scanner

Indicatorii tendinței de grupare frecvent folosiți sunt: media empirică, mediana, modul, definiți ca mai jos, pentru seria statistică (3.2.1) și, respectiv, variabila aleatoare empirică (3.2.5).

Media empirică sau **valoarea medie empirică** (*eng.* mean) sau speranța matematică a seriei considerate este un număr $\bar{x}$ definit astfel:

(a) atunci când nu se iau în considerare repetările unei valori (adică nu se utilizează frecvențele) este media aritmetică (simplă) a tuturor termenilor seriei

$$\bar{x} = \frac{x_1 + \ldots + x_N}{N} = \frac{\sum_{j=1}^{N} x_j}{N}; \qquad (3.2.12a)$$

(b) atunci când se iau în considerare repetările unei valori (adică se utilizează frecvențele absolute și relative) este media aritmetică ponderată a tuturor termenilor seriei, din care se exclud repetările

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{N} = \sum_{i=1}^{n} \upsilon_i x_i \,. \qquad (3.2.12b)$$

Este evident că media simplă se obține imediat din cea ponderată pentru $w_i = 1$ și, respectiv, $\upsilon_i = 1/N$. Cum era de așteptat, media ponderată coincide ca formulă cu media definită prin (3.1.4) pentru variabila aleatoare cu număr finit de valori (vezi subsecțiunea 3.1.2).

Media empirică, ca valoare, este întotdeauna cuprinsă între valoarea cea mai mică și cea mai mare, exprimând un nivel mediu, în care efectele valorilor extreme tind într-o oarecare măsură să se compenseze.

În general vorbind, media empirică este sensibilă la valorile extreme, astfel încât, dacă valorile din seria considerată sunt foarte "împrăștiate", media empirică poate capătă o valoare nereprezentativă pentru serie.

Mediana (*eng.* median) seriei considerate este un număr m_e definit astfel:

$$m_e = \begin{cases} x_{p+1} & \text{daca } N \text{ impar } (=2p+1), \\ (x_p + x_{p+1})/2 & \text{daca } N \text{ par } (=2p). \end{cases} \qquad (3.2.13)$$

Cu alte cuvinte, atunci când seria conține un număr impar de termeni, mediana este acel termen din șir plasat la mijlocul seriei (ca poziționare), la stânga și la dreapta medianei existând același număr de termeni. Când seria conține un număr par de termeni, se iau cei doi termeni din mijloc (ca poziționare) și li se calculează media aritmetică.

Dacă repartiția valorilor din serie este simetrică în raport cu mediana, media empirică este egală cu mediana.

Modul (*eng.* mode) sau dominanta seriei este acea valoare a caracteristicii măsurate care corespunde frecvenței relative maxime.

Indicatorii tendinței de împrăștiere frecvent folosiți sunt: dispersia empirică și abaterea pătratică medie empirică, definiți ca mai jos, pentru seria statistică (3.2.1).

Dispersia empirică sau **varianța** empirică (*eng.* variance) a seriei considerate este un număr $\overline{s_x}^2$ definit astfel:

Scanned with OKEN Scanner

(a) atunci când nu se iau în considerare repetările unei valori (adică nu se utilizează frecvențele) este media aritmetică (simplă) a tuturor abaterilor pătratice în raport cu media empirică

$$\overline{s_x}^2 = \frac{\sum_{j=1}^{N} (x_j - \bar{x})^2}{N}; \qquad (3.2.14a)$$

(b) atunci când se iau în considerare repetările unei valori (adică se utilizează frecvențele absolute și relative) este media aritmetică ponderată a tuturor abaterilor pătratice în raport cu media empirică, din care se exclud repetările

$$\overline{s_x}^2 = \frac{\sum_{i=1}^{n} w_i (x_i - \bar{x})^2}{N} = \sum_{i=1}^{n} \upsilon_i (x_i - \bar{x})^2 . \qquad (3.2.14b)$$

Dispersia empirică indică astfel gradul de împrăștiere a valorilor luate de variabila empirică în jurul valorii sale medii. O dispersie empirică de valoare mică reflectă o grupare importantă a valorilor din seria considerată, în jurul mediei empirice.

Trebuie precizat faptul că, în cazul selecțiilor de volum N relativ redus, se recomandă ca expresia dispersiei empirice dată de partea a) a definiției (adică 3.2.14a) să conțină la numitor valoarea N-1 în loc de N. Cum era de așteptat, dispersia empirică calculată ca medie ponderată în partea b) a definiției (adică 3.2.14b) coincide ca formulă cu dispersia definită prin (3.1.5) pentru variabila aleatoare cu număr finit de valori în subsecțiunea 3.1.2.

Din formula dispersiei empirice se observă imediat că unitatea de măsură a acesteia este pătratul unității de măsură pentru valorile variabilei aleatoare empirice.

În general vorbind, maniera de exprimare a dispersiei empirice prin raportare la media empirică este specifică unei clase largi de indicatori, denumiți **momente centrate** empirice de ordin q, definiți prin

$$(m_x)_q = \frac{1}{N} \sum_{j=1}^{N} (x_j - \bar{x})^q = \sum_{i=1}^{n} \upsilon_i (x_i - \bar{x})^q . \qquad (3.2.15)$$

Drept urmare, dispersia empirică este momentul centrat empiric de ordinul 2 al seriei statistice, sau, echivalent al variabilei aleatoare empirice.

Abaterea pătratică medie empirică sau **deviația standard** empirică (*eng.* standard deviation) se definește ca rădăcina pătrată a dispersiei empirice, conform formulei:

$$\overline{s_x} = \sqrt{\frac{\sum_{j=1}^{N} (x_j - \bar{x})^2}{N}} = \sqrt{\sum_{i=1}^{n} \upsilon_i (x_i - \bar{x})^2} . \qquad (3.2.16)$$

Unitatea de măsură coincide cu cea a valorilor variabilei aleatoare, ceea ce justifică introducerea deviației standard ca un criteriu suplimentar pe lângă dispersia empirică. Ca și în cazul dispersiei empirice, rămâne valabilă recomandarea privind utilizarea valorii N-1 în cazul selecțiilor de volum redus.

Pe lângă caracterizarea modului de grupare, respectiv împrăștiere, a valorilor unei variabile aleatoare empirice (echivalent unei serii statistice), variabilele empirice pot fi caracterizate și din punctul de vedere al formei, referindu-ne la **parametrii de formă** denumiți asimetrie și, respectiv, exces, definiți ca mai jos.

Scanned with OKEN Scanner

Asimetria (*eng.* skewness) seriei considerate este un număr a_x definit astfel:

$$a_x = (m_x)_3 / \overline{s_x}^3, \tag{3.2.17}$$

unde $(m_x)_3 = \frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^3 = \sum_{i=1}^{n}\upsilon_i(x_i - \bar{x})^3$ este momentul centrat empiric de ordinul 3 (obținut din relația generală (3.2.15) pentru $q = 3$), iar $\overline{s_x}$ (3.2.16) notează abaterea pătratică medie empirică discutată anterior.

Asimetria caracterizează modul în care valorile variabilei aleatoare empirice sunt plasate de o parte și de a alta a valorii medii $\bar{x}$. Pentru $a_x>0$ (asimetrie pozitivă) valorile sunt etalate preponderent la dreapta valorii medii $\bar{x}$, iar pentru $a_x<0$ (asimetrie negativă) valorile sunt etalate preponderent la stânga valorii medii $\bar{x}$.

Excesul (*eng.* kurtosis) seriei considerate este un număr e_x definit astfel:

$$e_x = (m_x)_4 / \overline{s_x}^4 - 3, \tag{3.2.18}$$

unde $(m_x)_4 = \frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^4 = \sum_{i=1}^{n}\upsilon_i(x_i - \bar{x})^4$ este momentul centrat empiric de ordinul 4 (obținut din relația generală (3.2.15) pentru $q = 4$), iar $\overline{s_x}$ (3.2.16) notează abaterea pătratică medie empirică discutată anterior.

Excesul caracterizează gradul de aplatizare sau de boltire a graficului frecvențelor, prin raportare la graficul unui echivalent de tip discret al legii de repartiție normală (de tip Gauss), cu aceeași medie $\bar{x}$ (3.2.12) și aceeași dispersie $\overline{s_x}^2$ (3.2.14). (Reamintim că repartiția normală este de tip continuu și, de aceea, este obligatorie echivalarea în valori discrete.) Pentru $e_x>0$ (exces leptocurtic) graficul frecvențelor are o alură boltită, iar pentru $e_x<0$ (exces platicurtic) graficul frecvențelor are o alură aplatizată.

Exemplul 3.2.2

Ne plasăm în contextul măsurătorilor de arii prezentate în exemplul 3.2.1, și al seriei statistice aferente. Media empirică calculată cu formula (3.2.12) este 371.1333. Mediana este 337, lăsând la stânga și, respectiv, la dreapta, câte 7 valori. Întrucât toate cele 15 arii măsurate sunt distincte, nu se pune problema evidențierii unui mod al seriei. Dispersia empirică și deviația standard empirică, calculate cu formulele (3.2.14) și, respectiv, (3.2.16), sunt 30756 și, respectiv, 175.3738. Dacă în formulele (3.2.14) și (3.2.16) utilizăm la numitor N-1=14, în loc de N=15 (deoarece seria statistică are puțini termeni), se va obține dispersia empirică 32953 și deviația standard empirică 181.5292. (Este evident că, în cazul unei serii statistice cu mult mai mulți termeni, adică N mult mai mare, diferențele dintre rezultatele obținute utilizând formulele cu N și, respectiv N-1 devin nesemnificative.)

Factorul de asimetrie calculat cu formula (3.2.17) are valoarea 1.1543, evidențiind etalarea preponderentă a valorilor la dreapta mediei empirice (asimetrie pozitivă). Acest aspect este vizibil pe reprezentările grafice de tip histogramă (din fig. 3.2.3(b), sau

Scanned with OKEN Scanner

fig. 3.2.4(b)) şi de tip poligon al frecvenţelor (din fig. 3.2.5(b)). Excesul, calculat cu formula (3.2.18), are valoarea 0.9512, evidenţiind un exces leptocurtic, adică o alură boltită. Acest aspect este vizibil pe reprezentările grafice de tip histogramă (din fig. 3.2.3(b), sau fig. 3.2.4(b)) şi de tip poligon al frecvenţelor (din fig. 3.2.5(b)).
Precizăm că toate caracteristicile evaluate anterior se referă la seria statistică, fără a lua în considerare gruparea pe clase a valorilor pe care am utilizat-o în exemplul 3.2.1, în scopul ilustrării tipurilor de reprezentări grafice. Totuşi, referirile la imaginile grafice de tip histogramă înseamnă implicit şi o grupare pe clase. Prin modificarea amplitudinii claselor, forma histogramei se modifică (după cum reiese comparând fig. 3.2.3(b) cu fig. 3.2.4(b)), dar, indiferent de modificare, tipul de asimetrie şi tipul de exces (evidenţiate prin calculul numeric asupra seriei) se vor regăsi vizual pe reprezentarea grafică.

3.2.5. DISTRIBUŢII TEORETICE ŞI DISTRIBUŢII EXPERIMENTALE

Uneori, în urma obţinerii unei serii statistice din măsurători, ne punem problema ca din datele experimentale (care definesc o **distribuţie empirică**) să deducem (cu o precizie cât mai bună) o **distribuţie teoretică** ce ar putea fi asociată celei empirice. Acesta este şi motivul pentru care, în secţiunea precedentă, am prezentat câteva noţiuni legate de variabilele aleatoare continue. Atât histograma cât şi poligonul frecvenţelor pot sugera vizual (în unele cazuri) alura densităţii de repartiţie a unei variabile aleatoare de tip continuu (cum ar fi o densitate de repartiţie uniformă – vezi fig. 3.1.2, sau o densitate de repartiţie normală – vezi fig. 3.1.3).

De exemplu, o histogramă cu aspectul general de clopot sugerează o lege normală (tip Gauss). Verificarea riguroasă a faptului că repartiţia empirică este de tip Gauss se abordează cu ajutorul unor teste speciale de verificare a ipotezelor statistice.

Prezentarea aparatului matematic aferent verificării ipotezelor statistice depăşeşte cadrul ce ni l-am propus pentru această lucrare, motiv pentru care cititorul interesat este invitat să consulte lucrări de specialitate (Mihoc, Micu, 1971, Rumşiski, 1974, Şabac et al., 1983).

De fapt, prelucrarea statistică cât mai eficientă a datelor experimentale mai presupune o serie de operaţii primare aplicate datelor, precum şi metode de construire a unei selecţii reprezentative ce nu sunt discutate în cartea de faţă. Neabordarea acestor problematici este deplin justificată de faptul că obiectivul central al acestei cărţi îl constituie tehnicile de realizare automată a măsurătorilor, şi nu utilizarea rezultatelor măsurătorilor în scopul generalizării lor pe baze statistice (ca estimaţii valide la nivelul unei întregi populaţii). Informaţiile din acest capitol sunt considerate însă necesare, deoarece evaluarea caracteristicilor cantitative ale tuturor entităţilor de un anumit tip existente într-un câmp prelucrat se confruntă, inevitabil, cu o anumită neuniformitate, care impune o **interpretare de natură statistică** (la nivel de eşantion sau selecţie).

Rămânând la nivelul intuitiv, vizual, al asocierii dintre o repartiţie empirică şi una teoretică, prezintă interes **reprezentarea grafică** a distribuţiei continue, comparativ cu cea empirică. În acest sens, pentru clase de amplitudini egale L, pentru fiecare clasă (interval) putem considera că distribuţia continuă are un echivalent discret, desemnând frecvenţa relativă cu care ar trebui luate valori din intervalul respectiv. Notând cu ω mijlocul unui interval generic, valoarea echivalentului discret al legii continue se va calcula cu formula

$$\text{echivalent_discret}(\omega) = L \times \text{distribuţia_continuă}(\omega). \qquad (3.2.19)$$

Scanned with OKEN Scanner

Egalitatea (3.2.19) rezultă din faptul că membrul ei drept aproximează (ca arie a unui dreptunghi având baza L şi înălţimea distribuţia_continuă(ω)) probabilitatea cu care legea continuă ia valori în intervalul generic din discuţie. Prin construirea grafică a echivalentului discret pentru legea normală, cu aceeaşi medie şi dispersie ca a variabilei aleatoare empirice, se pot aprecia intuitiv atât tipul de asimetrie, cât şi tipul de exces corespunzătoare seriei statistice.

Exemplul 3.2.3

Ne plasăm în contextul măsurătorilor de arii prezentate în exemplul 3.2.1, şi al seriei statistice aferente. Ne propunem să realizăm o comparaţie grafică între histogramele construite în exemplul 3.2.2, cu clase de amplitudine 100, respectiv, 50 (fig. 3.2.3(b), respectiv fig. 3.2.4(b)) şi echivalentul discret al legii normale cu aceeaşi medie şi dispersie ca a variabilei aleatoare empirice (adică, conform exemplului 3.2.2, medie 371.1333 şi dispersie 30756).

Echivalentul discret este reprezentat grafic cu simbolul "*" în fig. 3.2.8(a), pentru clase de amplitudine 100 (adică $L = 100$ în formula (3.2.19)), şi în fig. 3.2.8(b), pentru clase de amplitudine 50 (adică $L = 50$ în formula (3.2.19)). Conform fig. 3.2.8(a), obţinerea practică a unei repartiţii normale cu parametrii menţionaţi ar însemna ca rezultatele măsurătorilor să se plaseze cu următoarele frecvenţe relative: 0.1046, 0.1759, 0.2183, 0.2000, 0.1353, 0.0675, 0.0249, 0.0068, în intervalele [100, 200), [200, 300),....., respectiv, [800, 900].

Prin înjumătăţirea intervalelor, conform fig. 3.2.8(b), obţinerea practică a unei repartiţii normale cu parametrii menţionaţi ar însemna ca rezultatele măsurătorilor să se plaseze cu următoarele frecvenţe relative: 0.0438, 0.0613, 0.0795, 0.0955, 0.1064, 0.1099, 0.1052, 0.0933, 0.0767, 0.0585, 0.0413, 0.0271, 0.0164, 0.0092, 0.0048, 0.0023, în intervalele [100, 150), [150, 200),....., respectiv, [850, 900]. Din examinarea celor două desene din fig. 3.2.8, precum şi din calcularea parametrilor de formă (vezi exemplul 3.2.2), nu se poate lua o decizie categorică privind apropierea repartiţiei experimentale de una teoretică de tip Gauss. Evident că o atare decizie, rezultând preponderent din observaţii vizuale, este mai puţin fundamentată decât informaţiile ce s-ar obţine prin aplicarea unui test de normalitate.

O discuţie comparativă de aceeaşi factură se poate purta cu referire la echivalentul discret al repartiţiei continue de tip uniform pe intervalul [100, 900]. Echivalentul discret este reprezentat grafic cu simbolul "*" în fig. 3.2.9(a), pentru clase de amplitudine 100 (adică $L = 100$ în formula (3.2.19)), şi în fig. 3.2.9(b), pentru clase de amplitudine 50 (adică $L = 50$ în formula (3.2.19)). Conform fig. 3.2.9(a), obţinerea practică a unei repartiţii continue de tip uniform pe intervalul [100, 900] ar însemna ca rezultatele măsurătorilor să se plaseze cu frecvenţa relativă 0.1250 în fiecare din intervalele [100, 200), [200, 300),....., [800, 900]. Prin înjumătăţirea intervalelor, conform fig. 3.2.9(b), obţinerea practică a unei repartiţii continue de tip uniform pe intervalul [100, 900] ar însemna ca rezultatele măsurătorilor să se plaseze cu frecvenţa relativă 0.0625 în fiecare din intervalele [100, 150), [150, 200),....., respectiv, [850, 900]. Nici în acest caz nu se poate lua o decizie categorică privind apropierea repartiţiei experimentale de una teoretică de tip uniform.

Scanned with OKEN Scanner

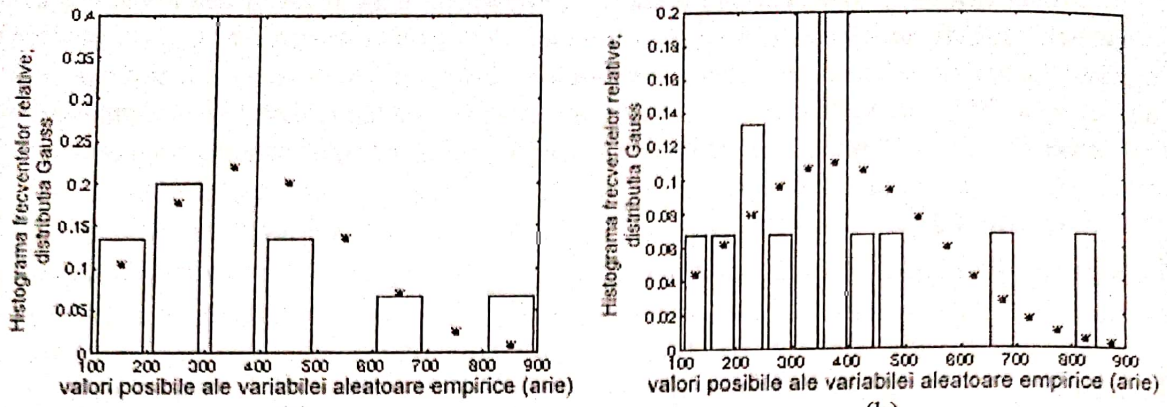


Fig. 3.2.8. Reprezentare grafică a echivalentului discret pentru legea normală de medie 371.1333 şi dispersie 30756: (a) pentru clase de amplitudine 100; (b) pentru clase de amplitudine 50.

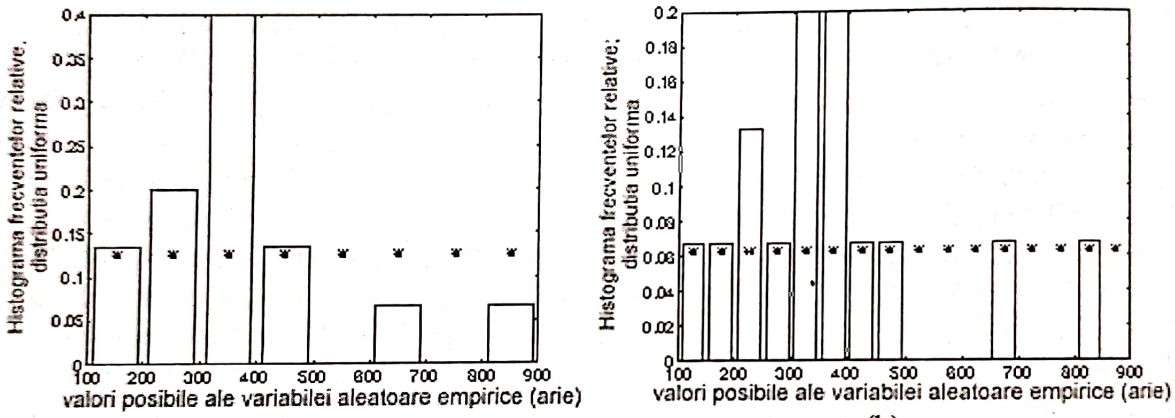


Fig. 3.2.9. Reprezentare grafică a echivalentului discret pentru o repartiție continuă de tip uniform pe intervalul [100-900]: (a) pentru clase de amplitudine 100; (b) pentru clase de amplitudine 50.

3.2.6. DEPENDENȚĂ STOHASTICĂ – METODA CORELAȚIEI ŞI A REGRESIEI

Există situații când, pentru entitățile ce fac obiectul măsurătorilor dintr-un câmp, se investighează **două caracteristici cantitative** φ şi ψ, despre care se doreşte a şti dacă depind una de cealaltă într-o anumită manieră. Astfel, în ordinea efectuării a N măsurători, pentru φ se obține secvența de valori:

$$x_1, x_2, \ldots, x_N, \tag{3.2.20}$$

iar pentru ψ se obține secvența de valori:

$$y_1, y_2, \ldots, y_N. \tag{3.2.21}$$

Posibila dependență dintre cele două variabile aleatoare empirice trebuie înțeleasă în sens **stohastic**, însemnând o interpretare probabilistică la nivelul ansamblului valorilor celor două secvențe (3.2.20), (3.2.21).

O primă constatare, de factură intuitivă, cu privire la posibila dependență dintre cele două variabile aleatoare empirice, se poate formula în urma construirii unei reprezentări grafice de tip **diagramă de dispersie** (*eng.* scattergram). Cu notațiile anterior considerate, se utilizează un reper de două axe ortogonale, pe axa absciselor măsurându-se valorile x_i, $i=1,\ldots,N$, iar pe axa ordonatelor, măsurându-se valorile y_i, $i=1,\ldots,N$. Pentru fiecare pereche $(x_i,y_i), i=1,\ldots,N$ se plasează câte un punct în diagrama de dispersie, astfel că în final aceasta va conține N puncte (unele dintre ele putând coincide – în cazul când există perechi (x_i,y_i) identice). În unele cazuri, examinarea vizuală a diagramei de dispersie poate sugera o dependență y funcție de x, care, la nivel global, se aseamănă cu reprezentări grafice binecunoscute (de tip liniar, pătratic). Atare concluzii, decurgând din intuiție și observație, pot fi formulate riguros în cadrul **metodei corelației**, completată de **metoda regresiei**. Fără a intra în totalitatea detaliilor matematice, vom prezenta pe scurt cele două metode, ca instrumente de explorare a dependenței stohastice.

Analiza **corelației** dintre variabilele aleatoare empirice φ și ψ se bazează pe semnificația coeficientului de corelație discutată în subsecțiunea 3.1.4. Conform formulei (3.1.18) și utilizând notațiile din secțiunea curentă (care se referă la variabile aleatoare empirice), **coeficientul de corelație** dintre φ și ψ poate fi scris astfel:

$$\rho_{xy}=\frac{\overline{(x-\bar{x})(y-\bar{y})}}{\overline{s_x}\,\overline{s_y}}. \qquad (3.2.22)$$

O valoare pozitivă a lui ρ_{xy} înseamnă că variabilele aleatoare empirice φ și ψ prezintă variații în același sens. O valoare negativă a lui ρ_{xy} înseamnă că variabilele aleatoare empirice φ și ψ prezintă variații în sensuri contrare.

Dacă modulul coeficientului de corelație $|\rho_{xy}|$ are o valoare apropiată de 1, înseamnă că între variabilele aleatoare empirice φ și ψ există o puternică dependență de factură liniară. Dacă modulul coeficientului de corelație $|\rho_{xy}|$ are o valoare apropiată de 0, înseamnă că dependența de factură liniară între variabilele aleatoare empirice φ și ψ este redusă (fără ca aceasta să excludă posibilitatea unei dependențe de factură neliniară). Dacă modulul coeficientului de corelație $|\rho_{xy}|$ are exact valoarea 1, înseamnă că între variabilele aleatoare empirice φ și ψ există o dependență liniară strictă.

Metoda regresiei exploatează informațiile specifice diagramei de dispersie (cu sau fără desenarea explicită a acesteia) și permite găsirea unei dependențe (liniare sau neliniare) a lui ψ ca funcție de φ, exprimată analitic, denumită **funcție de regresie**. Funcția de regresie are drept corespondent grafic în planul diagramei de dispersie o curbă, denumită **curbă de regresie**. În raport de informațiile furnizate de analiza corelației, persoana care examinează rezultatele statistice va utiliza un anumit tip de funcție de regresie.

Astfel, dacă modulul coeficientului de corelație $|\rho_{xy}|$ are o valoare apropiată de 1, se va utiliza o **funcție liniară** de forma:

$$\psi(\varphi)=A\varphi+B, \text{ sau, transcrisă în variabilele } x \text{ și } y,\ y(x)=Ax+B, \qquad (3.2.23)$$

cu coeficienții A și B urmând a fi determinați. Altminteri, se va face apel la o funcție neliniară. În alegerea funcției neliniare un rol important îl pot juca cunoștințele apriorice, cu caracter general, despre o posibilă dependență dintre caracteristicile φ și ψ. De asemenea,

Scanned with OKEN Scanner

interpretarea atentă a diagramei de dispersie poate oferi indicii valoroase pentru construirea unei **funcții neliniare**. De obicei, regresia neliniară utilizează funcții de tip parabolic:

$$\psi(\varphi) = A\varphi^2 + B\varphi + C \text{, sau, transcrisă în variabilele } x \text{ și } y,\ y(x) = Ax^2 + Bx + C, \quad (3.2.24)$$

funcții de tip hiperbolic:

$$\psi(\varphi) = A\frac{1}{\varphi} + B \text{, sau, transcrisă în variabilele } x \text{ și } y,\ y(x) = A\frac{1}{x} + B, \quad (3.2.25)$$

funcții de tip putere:

$$\psi(\varphi) = A\,B^{\varphi} \text{, sau, transcrisă în variabilele } x \text{ și } y,\ y(x) = A\,B^{x}. \quad (3.2.26)$$

În egalitățile de mai sus, coeficienții A, B, C au, inițial, valori necunoscute, iar odată ce a fost ales tipul funcției (și, implicit, al curbei), metoda regresiei are drept scop determinarea unor valori numerice optime. Criteriul de optimalitate îl constituie îndeplinirea **principiului "celor mai mici pătrate"**. Acest principiu, din punct de vedere geometric, presupune găsirea acelei curbe (din tipul ales) ce se află la cea mai mică distanță de toate punctele diagramei de dispersie. Minimizarea distanței de la curba de regresie la toate punctele diagramei de dispersie revine la a determina A, B, C astfel încât suma de pătrate

$$J = \sum_{i=1}^{N} (y_i - y(x_i))^2 \quad (3.2.27)$$

să ia cea mai mică valoare cu putință (de unde și denumirea principiului ca "cele mai mici pătrate"). În expresia lui J (3.2.27), valorile numerice x_i, y_i, $i = 1,\ldots,N$ sunt cunoscute din secvențele (3.2.20), (3.2.21), astfel încât J depinde numai de A, B, C, în raport cu care se realizează minimizarea. Valorile numerice determinate în urma minimizării pentru A, B, C, poartă denumirea de coeficienți de regresie.

După determinarea valorilor numerice ale coeficienților de regresie, funcția de regresie este complet cunoscută ca exprimare analitică și servește trasării curbei de regresie, care poate fi o dreaptă (în cazul funcției de tip liniar (3.2.23)), o parabolă (în cazul funcției de tip parabolic (3.2.24)), o hiperbolă (în cazul funcției de tip hiperbolic (3.2.25)) etc. Despre curba (funcția) de regresie se mai spune că realizează o fitare (*eng.* fitting) de tip liniar, parabolic, hiperbolic etc. a punctelor (x_i, y_i), $i = 1,\ldots,N$ obținute experimental.

În baza celor discutate, este evident că această curbă trece, în general, *printre* punctele din diagrama de dispersie și nu *prin* ele (excepție făcând situația când modulul coeficientului de corelație $|\rho_{xy}|$ are exact valoarea 1 și dreapta de dispersie trece *prin* toate punctele diagramei de dispersie). Am accentuat aspectul de mai sus, pentru a da și justificarea termenului de "regresie", în sensul că, prin intermediul curbei găsite, punctele experimentale de coordonate (x_i, y_i), $i = 1,\ldots,N$, nesituate pe curbă (adică nerespectând uniformitatea acesteia), capătă un corespondent "regresat" (adică tras înapoi pe curbă) de coordonate $(x_i, y(x_i))$, $i = 1,\ldots,N$.

În final, mai facem observația foarte importantă că, pentru orice tip de funcție de regresie, prin minimizarea sumei de pătrate J (3.2.27) se va intra în posesia unor valori numerice pentru coeficienții A, B, C. Aceasta nu înseamnă, însă, că respectivele valori numerice definesc cu adevărat o curbă (funcție) de regresie care exprimă o dependență între variabilele aleatoare empirice φ și ψ. Pentru a avea garanția unui asemenea rezultat, trebuie ca minimul sumei de pătrate J (3.2.27) să conducă la valoarea $\sqrt{J/N}$ relativ mică în comparație cu distanțele dintre punctele din diagrama de dispersie. Drept suport intuitiv pentru

Scanned with OKEN Scanner

semnificația acestei afirmații, sugerăm imaginea grafică a trei puncte situate în vârfurile unui triunghi echilateral; oricând vom putea duce o dreaptă situată la distanță minimă de cele trei puncte (prin minimizarea sumei de pătrate J (3.2.27)), dar nu vom putea spune că existența respectivei drepte ne îndreptățește a privi cele trei puncte ca aproximativ coliniare (echivalent, că vârfurile unui triunghi echilateral fac obiectul unei regresii liniare).

Exemplul 3.2.4

Ne plasăm în contextul exemplului 3.2.1. Pentru celulele reprezentate sub forma regiunilor albe considerate în fig. 3.2.1, realizăm o măsurare automată atât a ariilor, cât și a perimetrelor. În ordinea efectuării măsurătorilor, rezultatele sunt:

- arii: 695, 396, 459, 815, 234, 241, 372, 114, 417, 368, 190, 334, 319, 337, 276 [pixeli^2];
- perimetre: 122.3700, 116.1200, 100.8100, 158.7500, 74.8100, 71.4000, 83.7400, 43.9400, 105.5400, 83.0500, 59.8400, 78.0800, 91.4000, 85.4000, 100.3700 [pixeli].

Se obține diagrama de dispersie din fig. 3.2.10, în care pe abscisă se măsoară valorile ariilor, iar pe ordonată se măsoară valorile perimetrelor. Efectuarea calculelor în baza formulei (3.2.22) conduce la valoarea $\rho_{xy} = 0.9189$. Valoarea apropiată de 1 a factorului de corelație indică existența unei anumite dependențe liniare; pozitivitatea factorului de corelație arată că aria și perimetrul variază în același sens (fapt care, sub raport intuitiv, era de așteptat). Utilizând regresia liniară cu o funcție de tipul (3.2.23), obținem funcția de regresie $y(x) = 0.1402x + 39.6765$, căreia îi corespunde dreapta de regresie din fig. 3.2.11(a). Utilizând regresia parabolică cu o funcție de tipul (3.2.24), obținem funcția de regresie $y(x) = -0.0001x^2 + 0.1965x + 28.8295$, căreia îi corespunde parabola de regresie din fig. 3.2.11(b). Se remarcă faptul că termenul de gradul doi are un coeficient foarte mic în raport cu cel al termenului liniar, ceea ce face ca parabola să fie foarte aplatizată, apropiindu-se de o dreaptă (fapt ușor sesizabil în figură). Schimbând rolul variabilelor x și y, calitatea regresiei parabolice (adică x polinom de gradul 2 în y) nu se îmbunătățește. Astfel, putem concluziona că, pentru succesiunea de măsurători considerată în acest exemplu, dependența perimetru-arie este, în esență, de tip liniar. În final, subliniem faptul că validitatea funcțiilor de regresie obținute în acest exemplu este susținută de valoarea raportului $\sqrt{J}/N$ (10.5568 - pentru regresie liniară cu (3.2.23), 10.2963 - pentru regresie parabolică cu (3.2.24)), valoare relativ redusă în raport cu magnitudinea rezultatelor măsurătorilor.

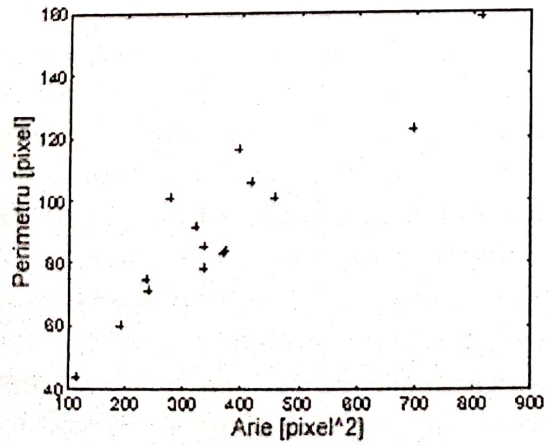


Fig. 3.2.10. Diagrama de dispersie aferentă perechilor (arie, perimetru) rezultate din măsurarea automată.

Scanned with OKEN Scanner

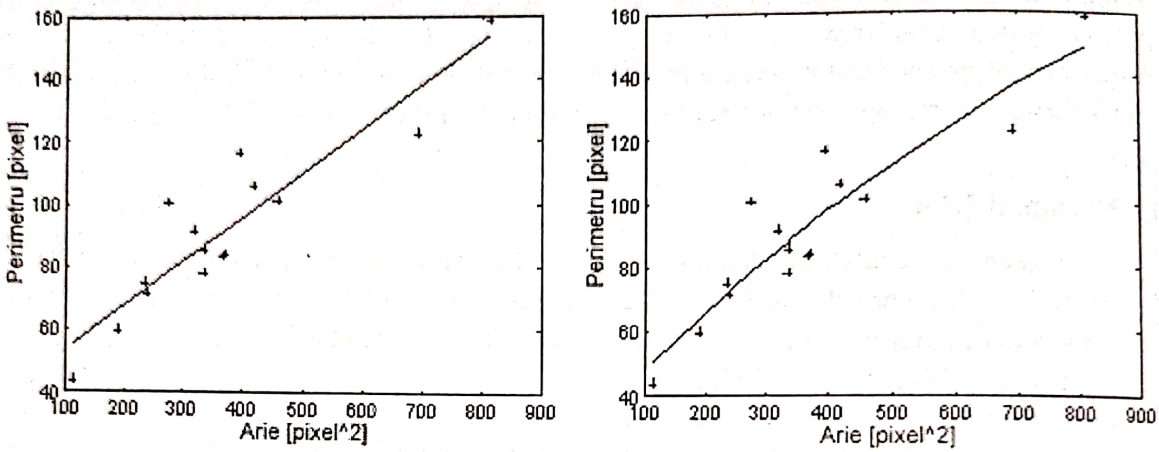


Fig. 3.2.11. Dependența perimetrului de arie stabilită prin regresie: (a) liniară; (b) parabolică.

Scanned with OKEN Scanner

Scanned with OKEN Scanner

CAPITOLUL 4

STUDII DE CAZ SEMNIFICATIVE

În medicină şi biologie, **procesarea de imagine** a făcut apel, iniţial (sfârşitul anilor `80, începutul anilor `90), la diverse produse program ce răspundeau strict **problematicii biomedicale** studiate. Spre deosebire de aceste realizări de pionierat (bazate în cea mai mare măsură pe programe elaborate artizanal (Connely, 1990, Ehricke et al., 1994, Hamilton, 1995)), sfârşitul anilor `90 a adus o evoluţie remarcabilă a mediilor software dedicate procesării de imagine, ca urmare a interesului manifestat de companii de prestigiu, care au elaborat produse profesionale, acoperind o arie largă de aplicaţii (de exemplu Zeiss KS400 (Kontron Elektronik GmbH, 2002, Carl Zeiss Vision GmbH, 2002), OPTIMAS (Media Cybernetics Inc., 2002), IMAGIC-5 (Image Science Software GmbH, 2002), NIH Image (National Institutes of Health – Research Services Branch, 2002), MATLAB (The MathWor, Inc., 2002)). În ciuda diversităţii lor, aceste produse recente prezintă *o serie de caracteristici comune, rezultând din fundamentul teoretic modern al procesării de imagine şi al vederii computerizate* (Russ, 1998, Seul, O`Gorman, Sammon, 2000, Pratt, 2001, Shapiro, Stockman, 2001, Gonzales, Woods, 2002).

În consecinţă, prin folosirea platformelor software sus-menţionate, pot fi realizate **măsurători complexe**, asigurând o înaltă acurateţe ce se datorează, pe de o parte, instrumentelor variate capabile să manipuleze imaginile la nivel de pixel şi, pe de altă parte, numeroaselor resurse disponibile pentru calcul matematic, statistică şi reprezentare grafică. Deşi literatura biomedicală conţine multe rapoarte privind semnificaţia rezultatelor în morfometria histologică şi histopatologică, *utilizarea concretă a unor astfel de platforme software nu este detaliată explicit, cu scopul deliberat de a proteja dreptul de proprietate intelectuală asupra tehnicilor de construire şi programare a algoritmilor* (Căruntu, 2002).

De aceea, avem convingerea că, pentru a oferi cititorilor un material cu eficienţă reală, prezenta lucrare trebuie să conţină informaţii edificatoare referitoare la modalitatea de operare propriu-zisă a unui anumit mediu software. Astfel, dacă în capitolele 2 şi 3 exemplele au avut un pronunţat caracter didactic, spre a ilustra pas cu pas elementele teoretice fundamentale pentru procesarea digitală şi interpretarea statistică a rezultatelor, acest capitol se axează pe *dezvoltarea şi exploatarea modulelor software ce răspund unor probleme uzuale de morfometrie microscopică*. În acest context, includerea ca secţiuni distincte a unor **studii de**

caz demonstrative, cu evidentă finalitate practică, generează o perspectivă globală asupra manipulării noțiunilor discutate în capitolele anterioare, prin integrarea lor în ansamblul unor aplicații reale. Algoritmii ce au stat la baza aplicațiilor software construite pentru aceste studii de caz au fost implementați ca funcții utilizator în mediul Zeiss KS400 (număr de licență 0400183), operațional la Departamentul de Histologie, U.M.F. Iași. Din acest motiv, în prima secțiune a capitolului curent, am considerat oportună o scurtă prezentare a facilităților cu aplicabilitate biomedicală puse la dispoziție de Zeiss KS400.

Procedurile computerizate ce deservesc studiile de caz ilustrează modul performant de efectuare a măsurătorilor dedicate morfologiei microscopice sub Zeiss KS400, iar discutarea lor este însoțită de textele sursă și scheme logice, furnizate sub formă de anexe.

Studiile de caz au fost formulate astfel încât principiile de realizare ale software-ului să poată fi ușor extrapolate pentru alte clase de probleme. Totodată, subliniem faptul că informațiile privind exploatarea Zeiss KS400 sunt suficient de detaliate spre a permite cititorului reproducerea tuturor experimentelor de analiză cantitativă fie sub Zeiss KS400, fie într-un alt mediu dedicat, prevăzut cu facilități similare (de exemplu OPTIMAS, IMAGIC-5, NIH Image, MATLAB).

4.1. SCURTĂ PREZENTARE A UTILIZĂRII MEDIULUI ZEISS KS400[1] ÎN DOMENIUL BIOMEDICAL

Multiple rapoarte de cercetare științifică informează asupra rezultatelor aplicării **mediului Zeiss KS400** în domeniul biomedical (Xu et al., 2001, Institute of Biologic and Life Science – Image Analysis Centre, 2002, Jarchow, 2002, Macovei et al., 2002, Tramontina et al., 2002), precum și în diverse alte tipuri de investigații (mineralogie, tehnologia materialelor, control de calitate) (Kindratenko, 1997, Chmelar, Sandvik, 2000, Malkusch, 2000), gama largă de aplicabilitate provenind din însăși filozofia de realizare a software-ului.

Software-ul Zeiss KS400 comunică cu utilizatorul prin intermediul unei **interfețe grafice**, organizate în jurul Ferestrei Principale (Main Window), echipată cu **meniuri** adecvate pentru procesarea imaginilor. Următoarea terminologie (Kontron Electronik GmbH, 1996) este utilizată pentru a caracteriza informațiile vizuale specifice, necesare măsurătorilor: **imagini** (aspecte fotografiate sau preluate prin videocameră, precum și rezultate ale procesării acestora) și **reprezentări grafice** (desene efectuate de utilizator într-un plan virtual, denumit **plan grafic**). În cadrul unei sesiuni de lucru, mediul Zeiss KS400 memorează imaginile cu care se operează, asociindu-le câte un număr din secvența naturală 1, 2, 3, …., . Toate imaginile pot fi vizualizate într-un format cu dimensiuni reduse ca un ansamblu denumit Gallery (*rom.* galerie).

Efectuarea măsurătorilor în vederea obținerii de informații numerice se bazează pe **concepte standard** din teoria și practica procesării imaginilor, ca de exemplu: **câmp**, **regiune** (într-un câmp), **punct** (într-o regiune), **desen** (raportat la o regiune). Sub raport cantitativ, fiecare din aceste concepte, la rândul său, poate fi caracterizat printr-o serie de **trăsături/proprietăți** de factură geometrică și, respectiv, densitometrică. Ca principiu fundamental de exploatare a rezultatelor măsurătorilor, accesul la acestea se face prin intermediul unei/unor **baze de date** proprii fiecărei aplicații; datele sunt scrise în bază/baze de către funcțiile de măsurare ale sistemului, și, respectiv, sunt extrase de către funcții destinate vizualizării, sub formă de **tabel** sau de **grafice**.

[1] Cu acordul Carl Zeiss Instruments SRL, distribuitor autorizat unic și exclusiv Zeiss pentru România.

Scanned with OKEN Scanner

Software-ul Zeiss KS400 operează cu următoarele **entităţi software**:

- **Fişiere**: ♦fişier de imagini (tif, gif, img, pcx etc.); ♦fişier program (mcr); ♦fişier cu informaţii referitoare la măsurători (ms); ♦fişier de baze de date (d00);
- **Ferestre de lucru** (după tipul de informaţii furnizate): ♦fereastră de imagini; ♦fereastră de mesaje; ♦fereastră de date rezultate din măsurători; ♦fereastră de program.

În cele ce urmează, prezentăm o scurtă trecere în revistă a meniurilor asociate ferestrei principale, enumerând acele funcţii de procesare a imaginilor cu directă utilitate în investigaţiile biomedicale (Căruntu, 2003 a). Parametrii acestor funcţii sunt setaţi prin intermediul unor **cutii de dialog**. Este evident că utilizatorul interesat în exploatarea efectivă a mediului Zeiss KS400 nu se poate limita la informaţiile de mai jos, consultarea manualului (Kontron Elektronik, 1996) fiind indispensabilă. (Menţionăm însă că manualul se adresează unor utilizatori experimentaţi în domeniul procesării de imagine, fără a lua în discuţie suportul teoretic general, care a fost prezentat în capitolele 2 şi 3 ale lucrării.)

- Meniul **Image** (*rom.* imagine) pune la dispoziţie o serie de funcţii elementare de manipulare a fişierelor de tip imagine şi a ferestrelor ce conţin imagini. Dintre aceste funcţii amintim: **Load** (*rom.* încărcare), **Save** (*rom.* salvare), **Display** (*rom.* afişare), **Copy** (*rom.* copiere), **Delete** (*rom.* ştergere) etc.
- Meniul **Graphic** (*rom.* grafic) pune la dispoziţie funcţii ce permit desenarea în planul grafic. Dintre aceste funcţii amintim: **Pixel, Vector, Rectangle** (*rom.* dreptunghi), **Polygon, Circle** (*rom.* cerc). Utilizarea planului grafic presupune suprapunerea sa peste o imagine, astfel încât desenarea să se poată raporta la elementele concrete ale imaginii.
- Meniul **Enhance** (*rom.* îmbunătăţire) pune la dispoziţie funcţii ce permit modificarea calităţii imaginilor (în sensul de creştere a calităţii). Dintre aceste funcţii amintim: **Contrast, Smooth** (*rom.* netezire), **Edges** (*rom.* muchii), **Shading** (*rom.* umbre).
- Meniul **Morphology** (*rom.* morfologie) pune la dispoziţie funcţii pentru prelucrarea imaginilor cu nuanţe de gri, de tip filtrare (cu diverşi operatori – Sobel, Roberts, Laplace etc.) şi operaţii de bază asupra unor mulţimi de pixeli (erodare, dilatare, deschidere, închidere etc.). Dintre aceste funcţii amintim: **Filters** (*rom.* filtre), **Basics** (*rom.* operaţii de bază).
- Meniul **Transform** (*rom.* transformare) pune la dispoziţie funcţii pentru transformări geometrice, de distanţă şi de culoare. Dintre aceste funcţii amintim: **Affine** (*rom.* afin), **Allign** (*rom.* aliniere), **Projection** (*rom.* proiecţie), **Distance** (*rom.* distanţă), **Fourier, Colour model** (*rom.* model de culoare).
- Meniul **Segment** (*rom.* segmentare) pune la dispoziţie funcţii pentru definiri de prag şi realizarea operaţiei de segmentare. Dintre aceste funcţii amintim: **Threshold** (*rom.* prag), **Adaptive** (*rom.* adaptare), **Region** (*rom.* regiune), **Contour** (*rom.* contur).
- Meniul **Binary** (*rom.* binar) pune la dispoziţie funcţii pentru prelucrarea imaginilor alb-negru prin operaţii logice la nivel de pixel, operaţii morfologice de tipul erodare, dilatare, deschidere, închidere asupra unor mulţimi de pixeli cu valori binare, adăugarea sau ştergerea unor mulţimi de pixeli în raport cu anumite criterii. Dintre aceste funcţii amintim: **Boolean** (*rom.* boolean), **Morphology** (*rom.* morfologie), **Utilities** (*rom.* utilităţi).
- Meniul **Measure** (*rom.* măsurare) pune la dispoziţie funcţii ce permit definirea obiectivelor şi condiţiilor pentru efectuarea măsurătorilor (cu parametri precizaţi în manieră implicită de un fişier de tip ms, sau în manieră explicită de către utilizator) şi realizarea efectivă a măsurătorilor (care depun rezultatele într-o bază/baze de date

Scanned with OKEN Scanner

proprie/proprii aplicației). Dintre aceste funcții amintim: **Set condition** (*rom.* setare condiții), **Set property** (*rom.* setare proprietăți), **Set features** (*rom.* setare trăsături), **Set frame** (*rom.* setare cadru), **Interactive** (*rom.* interactiv), **Automatic** (*rom.* automat), **Point** (*rom.* punct). (În terminologia proprie Zeiss KS400, efectuarea oricărei măsurători se realizează cu ajutorul unei imagini binare, denumită imagine-mască; pentru caracteristici geometrice, măsurarea se efectuează pe imaginea-mască, iar pentru caracteristici densitometrice, măsurarea se efectuează aplicând imaginea-mască peste imaginea inițială – color sau în nuanțe de gri.)

- Meniul **Digitiser** (*rom.* digitizor) pune la dispoziție (opțional) funcții de control pentru digitizorul de imagini.
- Meniul **Evaluate** (*rom.* evaluare) pune la dispoziție funcții de prezentare a informațiilor cantitative rezultate din măsurători, exprimate ca tabel sau exprimate grafic și însoțite de informații de natură statistică. Dintre aceste funcții amintim: **Data list** (*rom.* listă de date), **Data scattergram** (*rom.* scattergramă), **Data histogram** (*rom.* histogramă), **Cross tabulation** (*rom.* tabulare încrucișată) etc.
- Meniul Data (*rom.* date) pune la dispoziție funcții ce permit crearea unei baze de date, conținând informațiile numerice rezultate din măsurători și exploatarea acesteia. Dintre aceste funcții amintim: **New** (*rom.* nou), **First line** (*rom.* prima linie), **Get value** (*rom.* obținerea valorii), **Next line** (*rom.* următoarea linie) etc.

Observație. Principiile teoretice care stau la baza procesării de imagini digitale, indiferent de platforma software concretă, au fost detaliate în cadrul capitolului 2. Numeroasele exemple ilustrative din capitolul respectiv au fost în majoritate construite prin manipularea funcțiilor disponibile în Zeiss KS400. Suportul matematic necesar pentru interpretarea reprezentărilor grafice și a informațiilor statistice furnizate de Zeiss KS400 a fost dezvoltat în cadrul capitolului 3. În mod deliberat, graficele din capitolul 3 au fost realizate în mediul MATLAB, pentru a oferi cititorului și alte variante de plotare, în afara celor specifice Zeiss KS400 (la care se face apel în studiile de caz din acest capitol). Sub raport informațional, plotările din capitolul 3 sunt echivalente cu cele realizabile în Zeiss KS400, diferența fiind în exclusivitate grafică (tipuri de linii și simboluri de plotare, caractere, culori etc.).

Utilizarea Zeiss KS400 (Căruntu, 2003 a) în aplicații cu un grad de complexitate ridicat presupune folosirea eficientă a **limbajului KLIC** (Kontron Elektronik Language for Image Computing), care funcționează pe principiul de interpreter (execuția se realizează direct, linie cu linie, fără faze distincte de compilare și linkeditare). KLIC include operatorii aritmetici și logici, funcțiile matematice obișnuite, precum și structurile tipice de control if, while și for. Fiecărei funcții din meniurile comentate anterior îi corespunde o anumită instrucțiune (cuvânt rezervat în KLIC) cu o sintaxă prestabilită, care secvențiază parametrii selectabili prin cutia de dialog. Această facilitate permite scrierea completă a unui program în KLIC numai prin utilizarea cutiilor de dialog. Unitățile de program (funcțiile) scrise de către utilizator în KLIC sunt denumite **macrouri**. Experiența de utilizare a sistemului Zeiss KS400 creează *abilități de programare în KLIC care fac posibilă scrierea directă a instrucțiunilor*, fără a mai apela la cutiile de dialog, ceea ce, de fapt, reprezintă exploatarea profesionistă a software-ului (Căruntu, 2003 a). Textele sursă din Anexa II, ce însoțesc studiile de caz demonstrative din secțiunile următoare, constituie exemple elocvente de utilizare a programării în KLIC, particularizând prin probleme de morfologie microscopică. Din acest motiv, pentru realizarea unor obiective

Scanned with OKEN Scanner

similare, privind exploatarea funcțiilor Zeiss KS400 (de exemplu: crearea și accesarea bazelor de date, definirea trăsăturilor care fac obiectul măsurătorilor, construirea reprezentărilor grafice), au fost date mai multe soluții de programare în KLIC.

Individualitatea funcțiilor puse la dispoziție de Zeiss KS400 se referă doar la implementarea lor ca și programe, dar *suportul algoritmic folosit de acestea este în principiu comun tuturor produselor cu facilități similare* – de tipul celor referite anterior, în introducerea la acest capitol. Drept urmare, cititorul care se află în posesia unei platforme software de procesare a imaginilor poate utiliza exemplele din capitolele 2 și 3 drept exerciții de exploatare a platformei respective, parcurgând toate etapele care au fost descrise și comentate în acele exemple (eventual cu adaptările impuse de specificul fiecărui produs).

4. 2. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU STRUCTURI CELULARE. ILUSTRARE LA NIVELUL PULPEI DENTARE

Imaginea aleasă de la nivelul **pulpei dentare** (fig. 4.2.1) a surprins o zonă periferică cu **odontoblaste**, caracterizată printr-o densitate celulară medie. Pentru efectuarea investigației asistate de calculator, detaliată în cele ce urmează, am proiectat și implementat două macrouri denumite **ODO_P1** și **ODO_P2** (vezi textul sursă din Anexa II.1), scopul urmărit fiind măsurarea **ariei celulare** și, respectiv, **nucleare**, pentru 30 de odontoblaste (Căruntu, 2002).

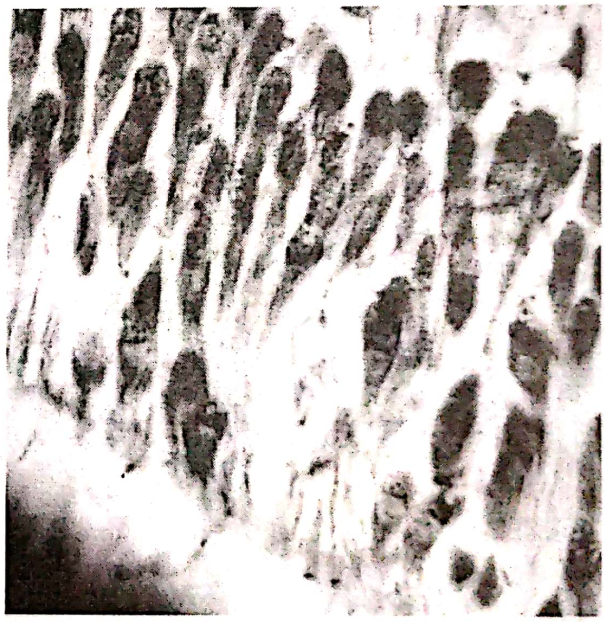


Fig. 4.2.1. Pulpă dentară – odontoblaste; imagine inițială pentru macroul ODO_P1.

4.2.1. DESCRIEREA PROCEDURII

Macroul **ODO_P1** a fost destinat marcării interactive a odontoblastelor, iar macroul **ODO_P2** a asigurat procesarea propriu-zisă, finalizată prin măsurători.

Selectarea interactivă a unui număr de 30 de odontoblaste (în **ODO_P1**) a fost realizată cu ajutorul editorului grafic (fig. 4.2.2). Inițial, au fost create două imagini binare complet negre, spre a servi drept suport pentru aplicarea planului grafic în care s-a realizat desenarea contururilor. Desenarea acestor contururi a fost făcută în două etape distincte: pentru celule și, respectiv, pentru nuclei. Prin aplicarea planului grafic cu contururile celulare

Scanned with OKEN Scanner

peste prima imagine binară a rezultat imaginea alb-negru din fig. 4.2.3, iar prin aplicarea planului grafic cu contururile nucleare peste a doua imagine binară a rezultat imaginea alb-negru din fig. 4.2.4. Concomitent, contururile desenate în planul grafic au fost suprapuse peste imaginea color inițială, obținându-se imaginea din fig. 4.2.5. Atât imaginea inițială, cât şi imaginile rezultate, au fost salvate în fişiere de imagine în format "tif", pentru a putea fi apelate ulterior, ca punct de plecare în macroul **ODO_P2**.

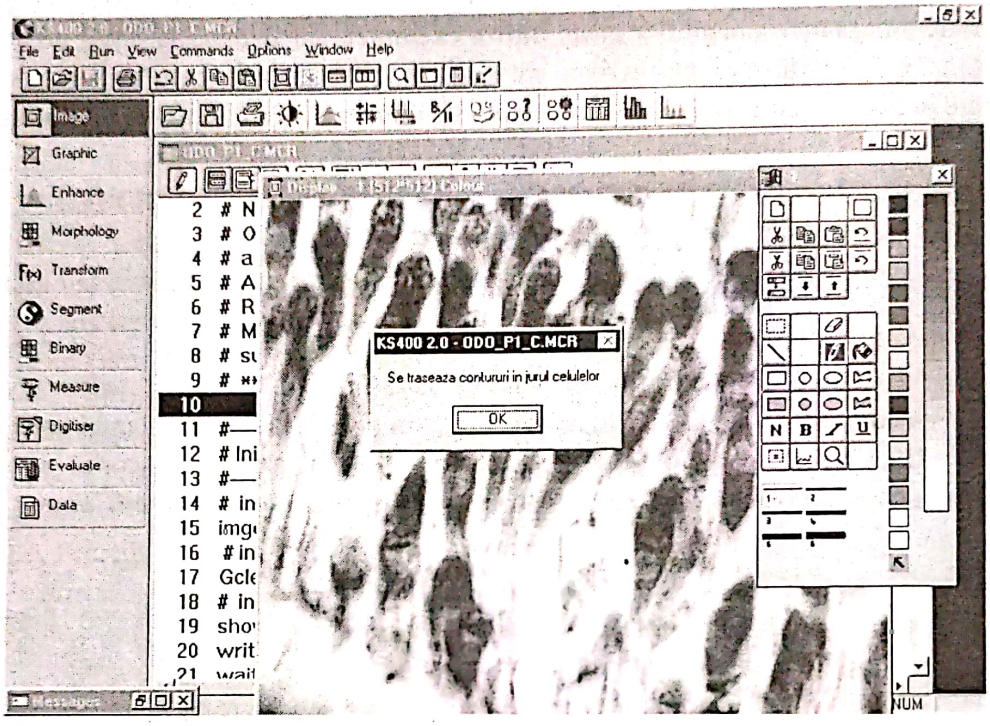


Fig. 4.2.2. Captură de ecran în timpul derulării macroului ODO_P1; editorul grafic lansat pentru trasarea interactivă de contururi.

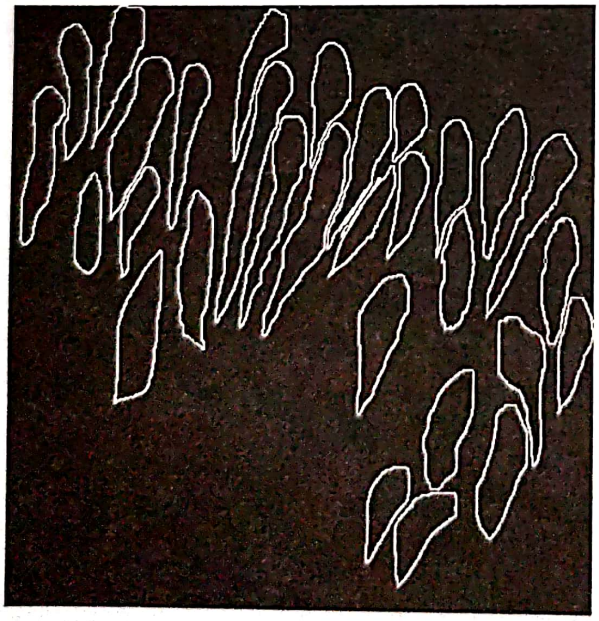


Fig. 4.2.3. Imagine binară rezultată în urma trasării interactive a contururilor celulare.

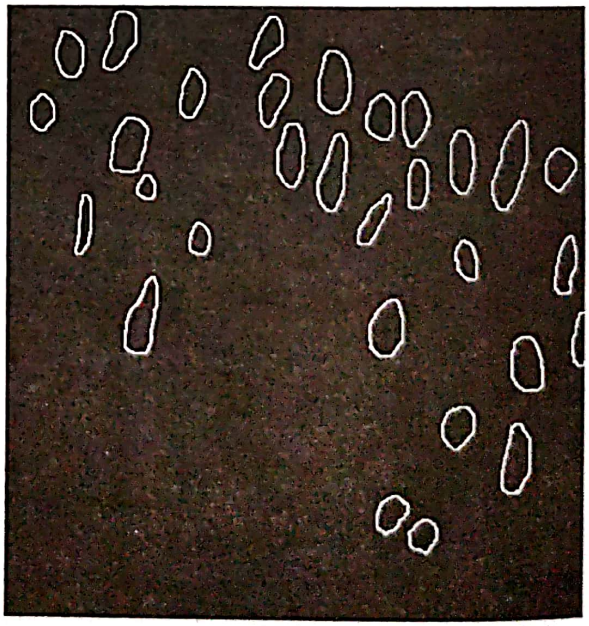


Fig. 4.2.4. Imagine binară rezultată în urma trasării interactive a contururilor nucleare.

Scanned with OKEN Scanner

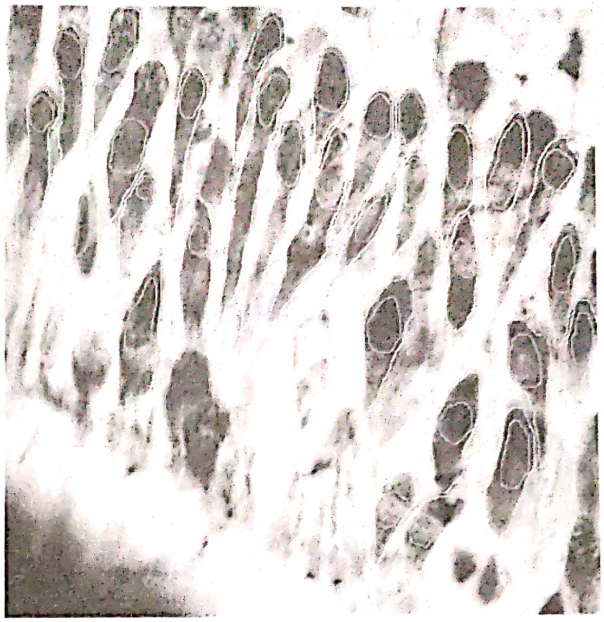


Fig. 4.2.5. Imaginea inițială peste care s-au aplicat contururile celulare și nucleare.

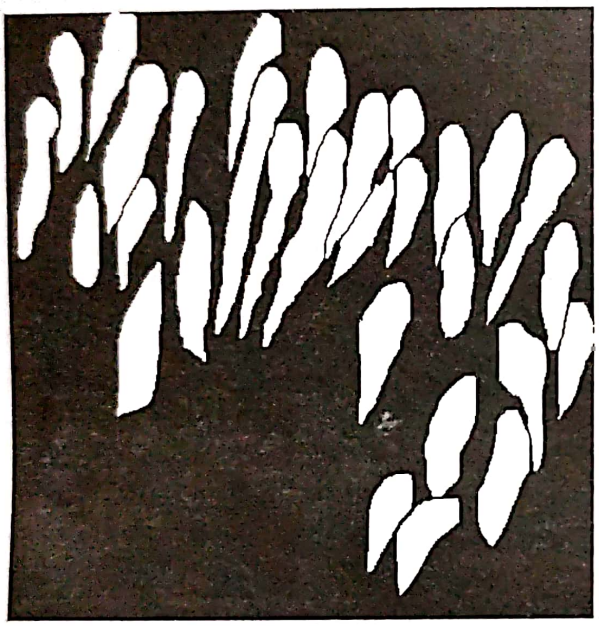


Fig. 4.2.6. Rezultatul operației de umplere a contururilor celulare.

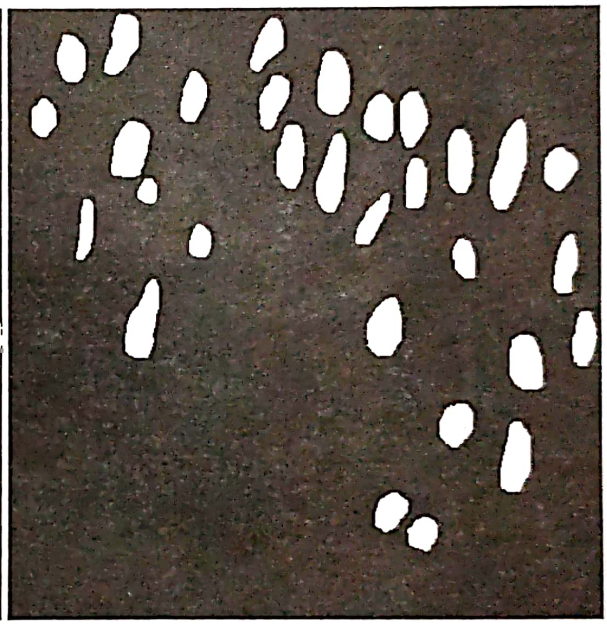


Fig. 4.2.7. Rezultatul operației de umplere a contururilor nucleare.

În macroul **ODO_P2**, pe imaginile din fig. 4.2.3 și, respectiv, fig. 4.2.4, se utilizează procedeul de umplere a găurilor, rezultând astfel două imagini binare, aferente celor 30 de contururi celulare (fig. 4.2.6) și, respectiv, nucleare (fig. 4.2.7).

Ulterior, aceste două imagini-mască au condus, prin operația booleană SAU-EXCLUSIV, la construirea imaginii-mască finale (fig. 4.2.8), utilizată pentru măsurători.

Măsurătorile au urmărit parametrii: **arie celulară** (variabila *CellArea*), **arie nucleară** (variabila *NcArea*) și **raport nucleo-citoplasmatic** (variabila *NcRatio*), care au fost definiți în debutul macroului **ODO_P2**. Informațiile numerice rezultate au fost memorate într-o bază de date nou creată, de unde au fost extrase și organizate de Zeiss KS400 într-o fereastră specială sub formă tabelară (fig. 4.2.9).

Scanned with OKEN Scanner

După cum se va vedea mai jos, informațiile din baza de date au fost de asemenea extrase în scopul realizării unor reprezentări grafice (de tip histogramă și diagramă de dispersie) și prelucrării statistice.

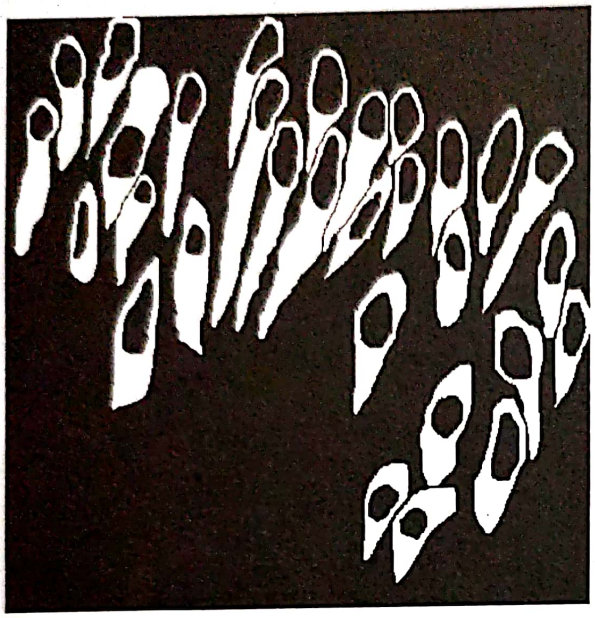


Fig. 4.2.8. Imagine-mască pentru măsurători.

Data List - Odonto (3,30)

File Edit Options Help

No	CellArea pixel2	NcArea pixel2	NcRatio %
1	2468.00	969.00	39.26
2	2268.00	755.00	33.29
3	2078.00	730.00	35.13
4	1356.00	814.00	60.03
5	2165.00	1165.00	53.81
6	3747.00	1153.00	30.77
7	2204.00	656.00	29.76
8	2687.00	507.00	18.87
9	2848.00	739.00	25.95
10	2241.00	934.00	41.68
11	2778.00	1481.00	53.31
12	1665.00	616.00	37.00
13	1592.00	481.00	30.21
14	1681.00	574.00	34.15
15	1481.00	258.00	17.42
16	3196.00	732.00	22.90
17	4133.00	745.00	18.03
18	3174.00	965.00	30.40
19	2188.00	491.00	22.44
20	4093.00	1102.00	26.92
21	2222.00	707.00	31.82
22	2872.00	393.00	13.68
23	3391.00	1195.00	35.24
24	1843.00	693.00	37.60
25	3011.00	1098.00	36.47
26	2802.00	1066.00	38.04
27	2833.00	779.00	27.50
28	3003.00	1117.00	37.20
29	1768.00	611.00	34.56
30	2244.00	510.00	22.73

Fig. 4.2.9. Tabel cu rezultatele numerice ale măsurătorilor realizate prin macroul ODO_P2.

Scanned with OKEN Scanner

4.2.2. INTERPRETAREA REZULTATELOR MORFOMETRICE

Determinarea, pentru fiecare din cele 30 de odontoblaste, a celor două tipuri de arii, a permis calcularea raportului dintre aria nucleului şi aria celulei (variabila *NcRatio*, care reprezintă un **descriptor bifazic** de morfometrie celulară).

Studiul statistic bazat pe valorile acestui tip de raport a fost reflectat într-o histogramă şi un rezumat cu mărimile matematice de interes (fig. 4.2.10), corespunzând la 14 clase echidistante, fiecare clasă acoperind un interval de 5 procente. Trasarea interactivă a contururilor introduce erori inevitabile în imaginea-mască (fig. 4.2.8), erori care, în condiţiile unui desen îngrijit, nu pot afecta raportul *NcRatio* cu mai mult de 1-2 procente. De aceea, prin alegerea valorii de 5 procente pentru fiecare clasă din reprezentarea grafică, s-a considerat că poate fi păstrată alura histogramei, chiar în prezenţa erorilor inerente datorate trasării interactive.

Asimetria este uşor în favoarea valorilor superioare mediei, fapt confirmat de factorul de asimetrie 0.629177. Comparând histograma cu curba Gauss caracterizată prin aceeaşi medie şi deviaţie standard (abatere), s-a constatat, atât grafic, cât şi analitic, forma moderat leptocurtică (exces = 0.487722), care înseamnă aglomerarea valorilor individuale în jurul valorii medii şi boltirea histogramei.

În ipoteza că eşantionul format din cele 30 de odontoblaste selectate este relevant, interpretarea datelor statistice revine la a afirma că, pentru majoritatea celulelor, aria nucleului reprezintă între 20% şi 45% din aria celulei (adică, aproximativ media ± deviaţia standard), cu o uşoară favorizare a valorilor cuprinse între 32.5% şi 45%.

La constatări echivalente din punct de vedere statistic am ajuns din analiza diagramei de dispersie arie nucleu versus arie celulă şi a rezumatului cu mărimile matematice de interes (fig. 4.2.11). De exemplu, unghiul cu vârful în origine, cuprins între semidreptele cu pantă 0.20 şi, respectiv, 0.45, conţine majoritatea punctelor, constituind analogul intervalului 20% - 45% din histogramă.

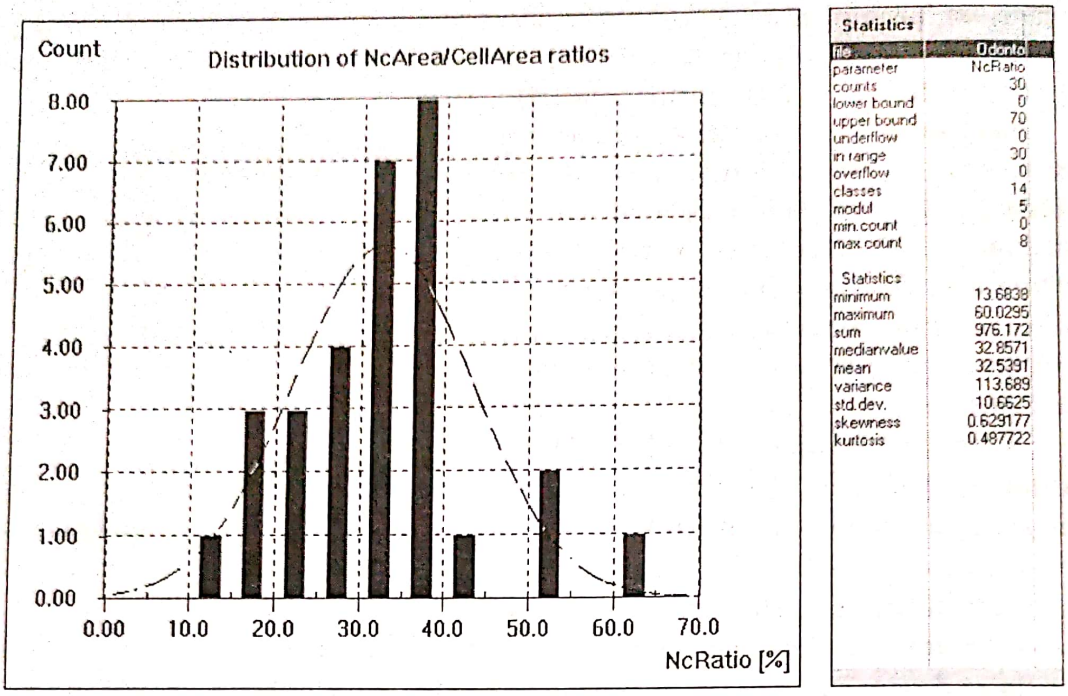
În plus, reprezentarea grafică de tip diagramă de dispersie a facilitat investigarea dependenţei dintre aria nucleilor şi aria celulelor, factorul de corelaţie obţinut (0.510507) nefiind edificator în sensul unei corelaţii liniare (care practic ar însemna o valoare subunitară, foarte apropiată de 1). Din acest motiv, fitarea polinomială de grad 1 (liniară), realizată de Zeiss KS400 asupra celor 30 de puncte din diagrama de dispersie, este mai mult formală, cu o eroare semnificativă, fapt imediat vizibil datorită numeroaselor puncte depărtate de dreapta de regresie. O remarcă de acelaşi tip se poate formula şi pentru fitarea polinomială de grad 2 (parabolică), ai cărei coeficienţi sunt de asemenea extraşi automat de Zeiss KS400.

Modul de repartiţie a ariilor celulare este ilustrat grafic în histograma de frecvenţe relative din fig. 4.2.12, care pune în evidenţă valoarea minimă de 1365 pixeli^2, valoarea maximă de 4133 pixeli^2 şi media de 2534.4 pixeli^2.

Modul de repartiţie a ariilor nucleare este ilustrat grafic în histograma de frecvenţe relative din fig. 4.2.13, din care rezultă valoarea minimă de 258 pixeli^2, valoarea maximă de 1481 pixeli^2 şi media de 801.2 pixeli^2.

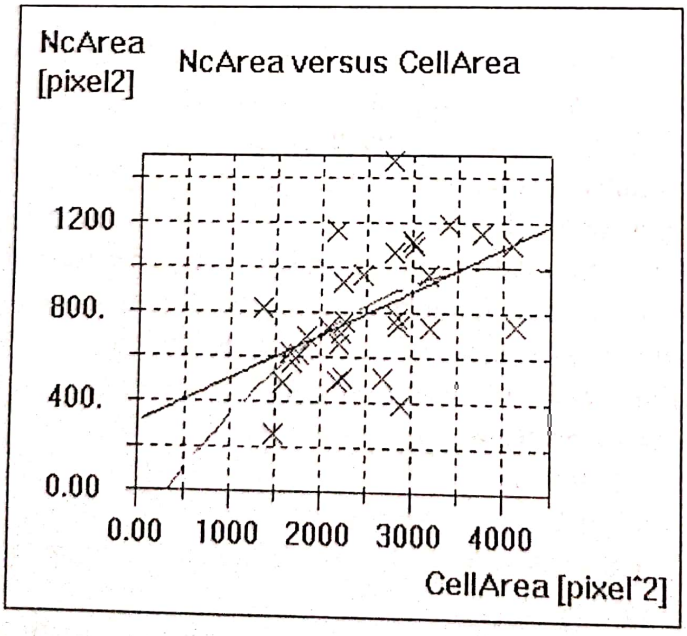
Confruntând alura grafică a histogramelor cu informaţiile de factură statistică, cititorul poate interpreta caracteristicile cantitative ale eşantionului de 30 de odontoblaste analizat, prin raportare la elementele standard discutate în capitolul 3, subsecţiunea 3.2.4.

Scanned with OKEN Scanner



Statistics	
file	Odonto
parameter	NcRatio
counts	30
lower bound	0
upper bound	70
underflow	0
in range	30
overflow	0
classes	14
modul	5
min.count	0
max.count	8
Statistics	
minimum	13.6838
maximum	60.0295
sum	976.172
median value	32.8571
mean	32.5391
variance	113.689
std.dev.	10.6625
skewness	0.629177
kurtosis	0.487722

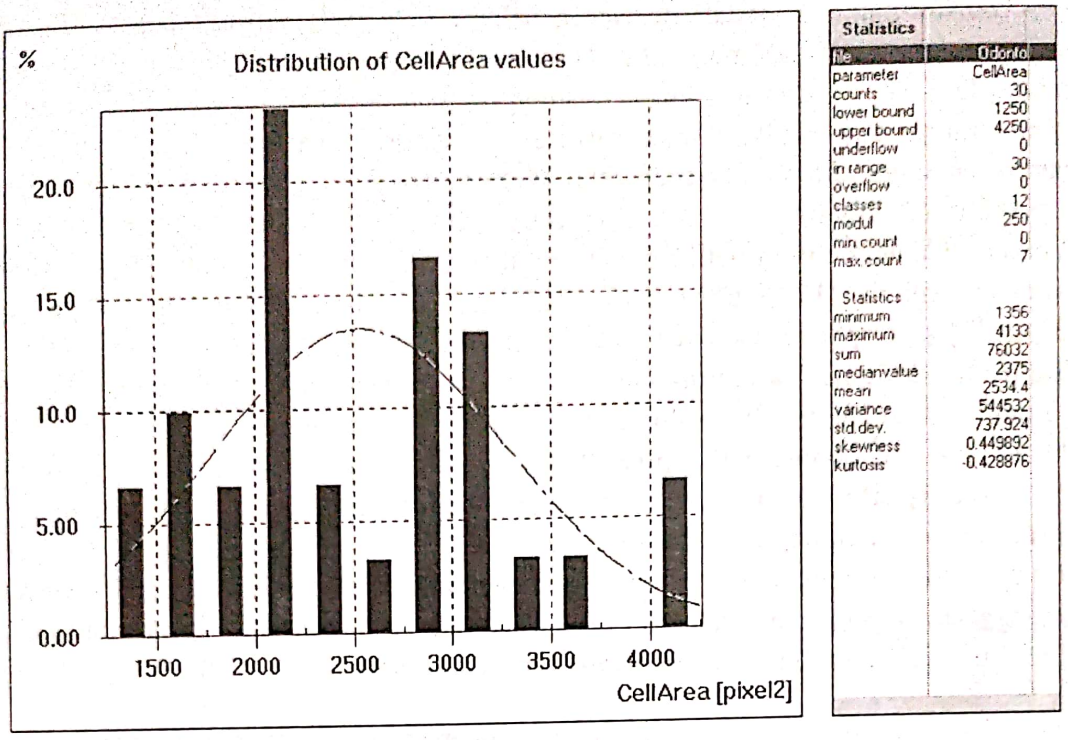
Fig. 4.2.10. Histogramă de frecvențe relative pentru valorile (subunitare) ale raportului arie nucleară / arie celulară, însoțită de informațiile statistice standard specifice Zeiss KS400.



Statistics	
file X	Odonto
parameter X	CellArea
file Y	Odonto
parameter Y	NcArea
counts	30
lower bnd X	0
upper bnd X	4500
lower bnd Y	0
upper bnd Y	1500
in range	30
out range	0
Statistics	
covariance	106312
correlation	0.510507
polynom 1	306.395
	0.195235
polynom 2	-180.037
	0.585536
	-7.23422e-005

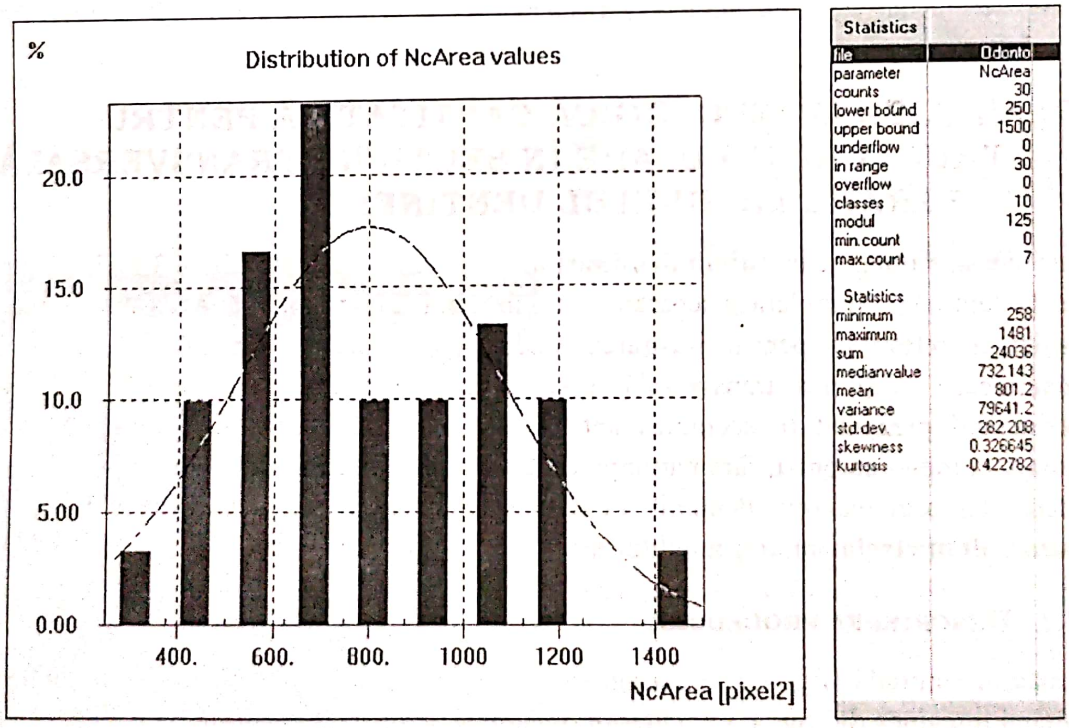
Fig. 4.2.11. Diagramă de dispersie reprezentând relația dintre aria nucleară și aria celulară, însoțită de informațiile matematice standard specifice Zeiss KS400.

Scanned with OKEN Scanner



Statistics	
file	Odontol
parameter	CellArea
counts	30
lower bound	1250
upper bound	4250
underflow	0
in range	30
overflow	0
classes	12
modul	250
min.count	0
max.count	7
Statistics	
minimum	1356
maximum	4133
sum	76032
medianvalue	2375
mean	2534.4
variance	544532
std.dev.	737.924
skewness	0.449892
kurtosis	-0.428876

Fig. 4.2.12. Histogramă de frecvențe relative pentru valorile ariei celulare, însoțită de informațiile statistice standard specifice Zeiss KS400.



Statistics	
file	Odontol
parameter	NcArea
counts	30
lower bound	250
upper bound	1500
underflow	0
in range	30
overflow	0
classes	10
modul	125
min.count	0
max.count	7
Statistics	
minimum	258
maximum	1481
sum	24036
medianvalue	732.143
mean	801.2
variance	79641.2
std.dev.	282.208
skewness	0.326645
kurtosis	-0.422782

Fig. 4.2.13. Histogramă de frecvențe relative pentru valorile ariei nucleare, însoțită de informațiile statistice standard specifice Zeiss KS400.

Scanned with OKEN Scanner

Transformarea în unităţi metrice a valorilor în pixeli obţinute pentru aria celulară şi aria nucleară se realizează pornind de la faptul că imaginea iniţială (fig. 4.2.1) este un pătrat cu latura de 8.68 cm, transpus într-o imagine digitală de 512x512 pixeli. Pe de altă parte, imaginea iniţială este rezultatul unei măriri de 400 de ori a preparatului microscopic, ceea ce înseamnă că unui micron din preparatul microscopic îi corespund 2.35 pixeli (deoarece 512 pixeli/(8.68 cm/400) = 2.35 pixeli/μm). Drept urmare, unui micron^2 îi corespund 5.52 pixeli^2. Cu aceste informaţii, conversia din pixeli în microni a ariilor celulare conduce la valoarea minimă de 245.65 μm^2 (pentru 1365 pixeli^2), valoarea maximă de 749 μm^2 (pentru 4133 pixeli^2) şi, respectiv, valoarea medie de 459.13 μm^2 (pentru 2534.4 pixeli^2). Procedând analog pentru ariile nucleare, prin conversia din pixeli în microni se obţin: valoarea minimă 46.73 μm^2 (pentru 258 pixeli^2), valoarea maximă 268.29 μm^2 (pentru 1481 pixeli^2) şi valoarea medie 145.14 μm^2 (pentru 801.2 pixeli^2). Aria celulară medie (459.13 μm^2), coroborată cu forma celulei (piriformă), permite să afirmăm că odontoblastele măsurate pot fi, în medie, încadrate într-un dreptunghi cu latura mare de aproximativ 50 μm (şi latura mică de circa 6-7 ori mai redusă). Valoarea de 50 μm constituie o bună confirmare a acurateţei măsurătorilor, deoarece apare raportată în literatura de specialitate ca dimensiune tipică (corespunzătoare înălţimii) pentru odontoblast (Avery, 1994, Ten Cate, 1998, Căruntu, 2001 a).

Observaţie. Acest exemplu de imagine microscopică a fost anume ales pentru a ilustra exploatarea Zeiss KS400 atunci când se operează cu două (sau mai multe) tipuri de entităţi structurale, apropiate ca şi culoare. În limbajul de procesare a imaginilor, aceasta înseamnă două (sau mai multe) tipuri de regiuni, care, datorită culorilor asemănătoare, nu pot fi individualizate automat, indiferent de nuanţările realizabile prin diverse tehnici de filtrare. Drept urmare, atare situaţii necesită lansarea modului interactiv de lucru.

4.3. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU STRUCTURI TUBULARE ÎN SECŢIUNE TRANSVERSALĂ. ILUSTRARE LA NIVELUL DENTINEI

A fost aleasă o imagine cu **tubuli dentinari** în secţiune transversală (fig. 4.3.1), înconjuraţi de inele de dentină peritubulară şi separaţi de arii de dentină intertubulară. Această imagine a fost considerată relevantă pentru ilustrarea studiului de caz propus, vizând trăsăturile ce caracterizează secţiunea transversală a structurilor tubulare. Spre deosebire de studiul experimental prezentat în secţiunea anterioară, am urmărit elaborarea unui procedeu de măsurare complet automat, fără participarea interactivă a utilizatorului. Acest procedeu a fost implementat prin macroul denumit **TUBDEN** (vezi textul sursă din Anexa II.2), dedicat măsurării **diametrelor** tuturor tubulilor dentinari din imaginea analizată (Căruntu, 2002).

4.3.1. DESCRIEREA PROCEDURII

Pe imaginea iniţială (fig. 4.3.1) s-a aplicat o tehnică adecvată de delimitare a contururilor, rezultând imaginea din fig. 4.3.2. Ulterior, imaginea rezultată a fost supusă unei segmentări de tip prag, care a condus la imaginea din fig. 4.3.3. Pe această imagine s-a practicat o operaţie morfologică de închidere, urmată de îndepărtarea regiunilor albe cu arii mai mici de 50 pixeli^2. Astfel, s-a obţinut imaginea din fig. 4.3.4, ce serveşte drept mască de tip binar pentru măsurătorile de efectuat sub Zeiss KS400.

Scanned with OKEN Scanner

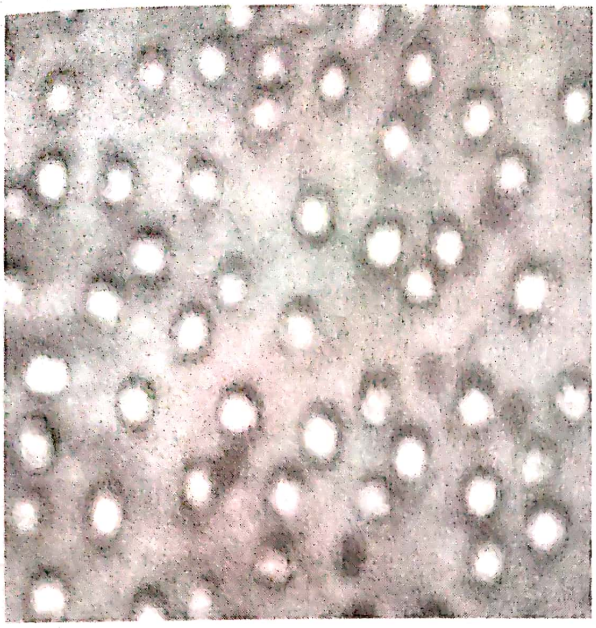


Fig. 4.3.1. Dentină – tubuli dentinari în secțiune transversală; imagine inițială pentru macroul TUBDEN.

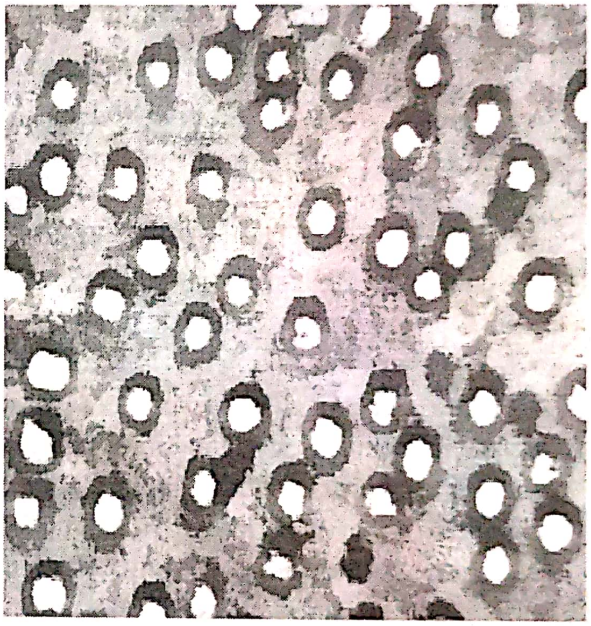


Fig. 4.3.2. Rezultatul aplicării operației de delimitare a contururilor pe imaginea din fig. 4.3.1.

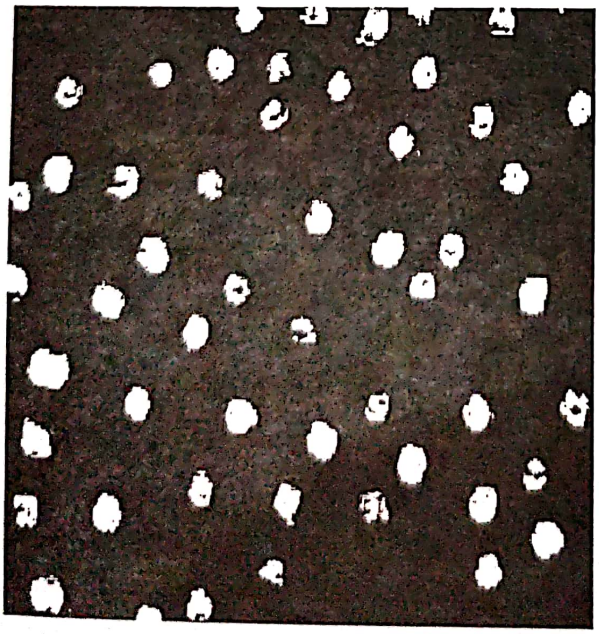


Fig. 4.3.3. Imagine binară rezultată în urma segmentării imaginii din fig. 4.3.2.

Scanned with OKEN Scanner

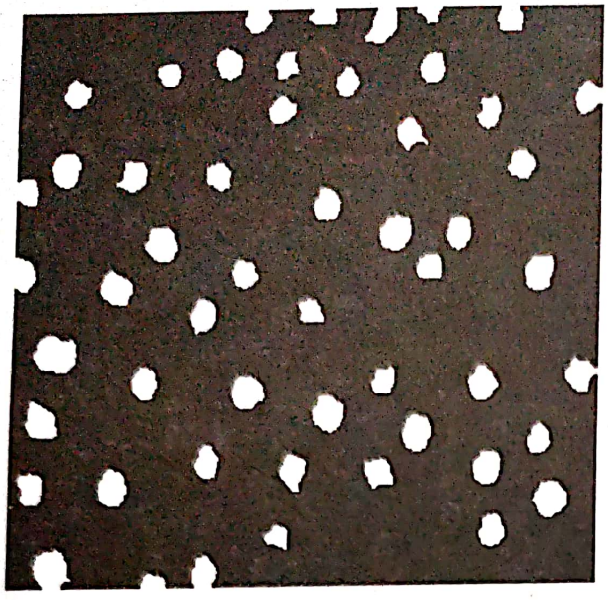


Fig. 4.3.4. Imagine-mască pentru măsurători, rezultată ca urmare a unei succesiuni de operații: închidere și îndepărtarea ariilor nesemnificative.

În alegerea parametrilor pentru măsurători, s-a pornit de la ideea că, în termeni histologici de factură calitativă, secțiunea transversală a tubulilor dentinari este considerată circulară. Sub raport strict geometric, însă, imaginile histologice dovedesc inevitabile abateri de la un contur circular, fapt datorat procesului de dentinogeneză în sine, și accentuat de artefactele inerente prelucrării prin demineralizare. În consecință, pentru exprimarea cantitativă a dimensiunii tubulilor am stabilit drept criteriu geometric riguros conceptul de **diametru echivalent** (variabila *DCIRCLE*) care, pentru un tubul dentinar oarecare, se definește drept diametrul discului având aceeași arie cu cea a tubulului considerat. Utilizarea acestui criteriu a necesitat verificarea, în termeni cantitativi, a ipotezei de circularitate asupra secțiunii transversale a tubulilor. Astfel, pe lângă măsurarea diametrului echivalent pentru regiunile delimitate de conturul real al tubulului, s-a avut în vedere calcularea a doi descriptori, și anume un **factor de convexitate** și un **factor de formă**. Factorul de convexitate a fost exprimat prin variabila:

$$AREA2AREAC = \frac{AREA}{AREAC},$$

și a făcut posibilă aprecierea gradului de apropiere a conturului real de un contur convex. Factorul de formă a fost exprimat prin variabila:

$$FCIRCLE = 4\pi \frac{AREA}{PERIMCROFT^2},$$

și a permis aprecierea gradului de apropiere dintre conturul real al tubulului și conturul ideal, de tip cerc.

Notațiile de mai sus au următoarele semnificații: *AREA* – aria regiunii delimitate de conturul real, *AREAC* – aria suprafeței ce reprezintă închiderea convexă a regiunii delimitate de conturul real, *PERIMCROFT* – perimetrul Crofton corespunzător regiunii delimitate de conturul real, fiind nume rezervate pentru Zeiss KS400. Ambii factori considerați sunt subunitari; condiția *AREA2AREAC*=1 se îndeplinește numai în cazul regiunilor convexe, iar condiția *FCIRCLE*=1 este satisfăcută numai pentru regiuni delimitate de contururi circulare.

Scanned with OKEN Scanner

Observaţie. Examinând expresia factorului de formă *FCIRCLE* se constată că este identică cu formula (2.6.5) furnizată în capitolul 2, dacă perimetrul se calculează cu formula Crofton (de unde şi numele variabilei *PERIMCROFT*). În cazul structurilor noastre de tip cvasi-circular, rezultatul evaluării numerice a perimetrului prin intermediul acestei formule nu diferă mult de cel ce se poate obţine cu ajutorul formulei (2.6.3), dar principiul de evaluare este cu totul altul. Modalitatea de calcul a perimetrului în sensul lui Crofton se bazează pe numărul de interceptări (definite ca numărul de intrări într-o regiune de-a lungul unei direcţii date) (Weisstein, 2002). S-a preferat utilizarea formulei lui Crofton deoarece, în primul rând, unele referinţe bibliografice o recomandă mai precisă decât (2.6.3) şi, în al doilea rând, ea există implementată ca funcţie Zeiss KS400. Mai mult, Zeiss KS400 are implementată o funcţie pentru calculul factorului de formă *FCIRCLE*, bazată pe evaluarea perimetrului prin metoda Crofton. Deşi nu este cazul imaginilor din exemplul curent, precizăm că unul din avantajele perimetrului calculat în sens Crofton îl constituie posibilitatea ca, în cazul pixelilor nepătraţi, să se opereze cu o precizie cert superioară algoritmului (2.6.3).

Ca urmare a măsurătorilor automate, informaţiile numerice rezultate au fost memorate într-o bază de date nou creată, de unde au fost extrase şi afişate de Zeiss KS400 într-o fereastră specială sub formă tabelară, din care fig. 4.3.5 reproduce primele 34 de linii din totalul de 53 (corespunzând numărului de tubuli prezenţi în imaginea din fig. 4.3.1). Ulterior, după cum se va vedea mai jos, informaţiile din baza de date au fost de asemenea extrase în scopul realizării unor reprezentări grafice (diagramă de dispersie, reprezentare tridimensională şi histogramă) şi prelucrării statistice.

Data List - tub_den (5,53)

File Edit Options Help

No	AREA pixel^2	AREAC pixel^2	AREA2AREAC	FCIRCLE	DCIRCLE pixel
1	294.00	310.14	0.95	0.86	19.35
2	213.00	260.15	0.82	0.68	16.47
3	377.00	406.38	0.93	0.92	21.91
4	424.00	448.95	0.94	0.90	23.23
5	611.00	684.55	0.89	0.78	27.89
6	476.00	513.00	0.93	0.91	24.62
7	329.00	389.68	0.84	0.76	20.47
8	309.00	329.37	0.94	0.95	19.84
9	440.00	471.52	0.93	0.92	23.67
10	346.00	382.73	0.90	0.88	20.99
11	398.00	442.58	0.90	0.90	22.51
12	494.00	544.42	0.91	0.87	25.08
13	395.00	430.79	0.92	0.93	22.43
14	413.00	484.67	0.85	0.77	22.93
15	428.00	541.06	0.79	0.68	23.34
16	432.00	476.96	0.91	0.89	23.45
17	590.00	627.41	0.94	0.94	27.41
18	476.00	557.89	0.85	0.79	24.62
19	374.00	410.24	0.91	0.89	21.82
20	374.00	398.21	0.94	0.91	21.82
21	473.00	519.03	0.91	0.88	24.54
22	632.00	679.44	0.93	0.93	28.37
23	452.00	486.36	0.93	0.91	23.99
24	612.00	660.33	0.93	0.92	27.91
25	371.00	395.24	0.94	0.93	21.73
26	446.00	474.19	0.94	0.85	23.83
27	372.00	416.43	0.89	0.90	21.76
28	587.00	619.77	0.95	0.92	27.34
29	573.00	625.94	0.92	0.91	27.01
30	304.00	341.71	0.89	0.81	19.67
31	450.00	491.41	0.92	0.88	23.94
32	911.00	990.83	0.92	0.91	34.06
33	331.00	361.82	0.91	0.88	20.53
34	521.00	620.09	0.84	0.78	25.76

datalist "tub_den",0,0 <0> 78 ms NUM

Fig. 4.3.5. Tabel cu rezultatele numerice ale măsurătorilor realizate prin macroul TUBDEN.

Scanned with OKEN Scanner

4.3.2. INTERPRETAREA REZULTATELOR MORFOMETRICE

Reprezentarea grafică de tip diagramă de dispersie (fig. 4.3.6) a pus în evidență faptul că majoritatea regiunilor măsurate se caracterizează prin valori situate în domeniul *AREA2AREAC*≥0.85, respectiv *FCIRCLE*≥0.80, dependența dintre cele două variabile (justificată prin valoarea 0.86 a factorului de corelație) putând fi aproximată drept o dependență liniară. Reprezentarea tridimensională (fig. 4.3.7), care operează cu clase echidistante (ce acoperă un interval de 5 procente atât pentru factorul de convexitate, cât și pentru factorul de formă), arată că în domeniul anterior referit se încadrează un număr de 44 de regiuni din totalul de 53 care au făcut obiectul măsurătorilor. Astfel, am obținut *confirmarea ipotezei calitative privind circularitatea contururilor* pentru regiunile investigate și, implicit, **diametrul echivalent** s-a constituit într-un instrument geometric adecvat pentru studierea dimensiunilor tubulilor dentinari. Histograma (fig. 4.3.8) organizată pe 10 clase echidistante, a câte 2 pixeli fiecare, a evidențiat concentrarea în jurul valorii medii de 24 de pixeli pentru diametrul echivalent. Conform informațiilor statistice furnizate de Zeiss KS400, repartiția empirică reprezentată de histogramă etalează o asimetrie ușor pozitivă (factor de asimetrie 0.271876) și o alură leptocurtică (exces = 0.55086). Aceste constatări au putut fi confirmate și prin compararea histogramei cu echivalentul discret al distribuției de tip gaussian cu aceeași medie și deviație standard (suprapusă cu linie punctată peste histograma din fig. 4.3.8).

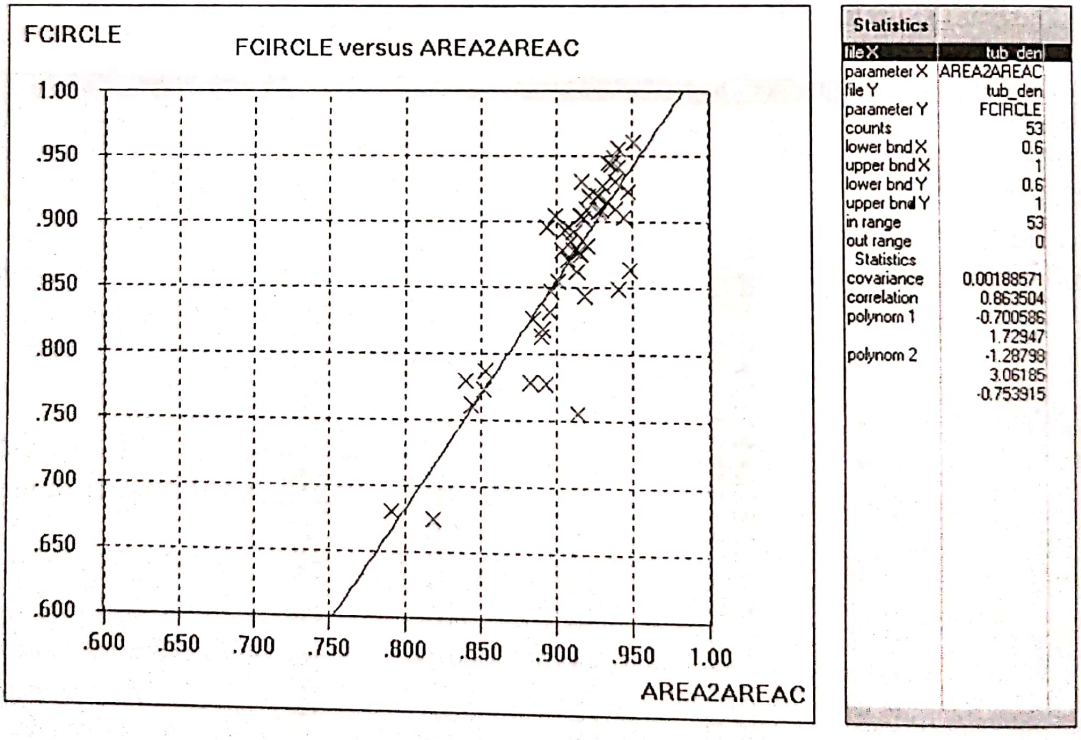


Fig. 4.3.6. Diagramă de dispersie reprezentând relația dintre factorul de convexitate și factorul de formă, însoțită de informațiile matematice standard specifice Zeiss KS400.

Scanned with OKEN Scanner

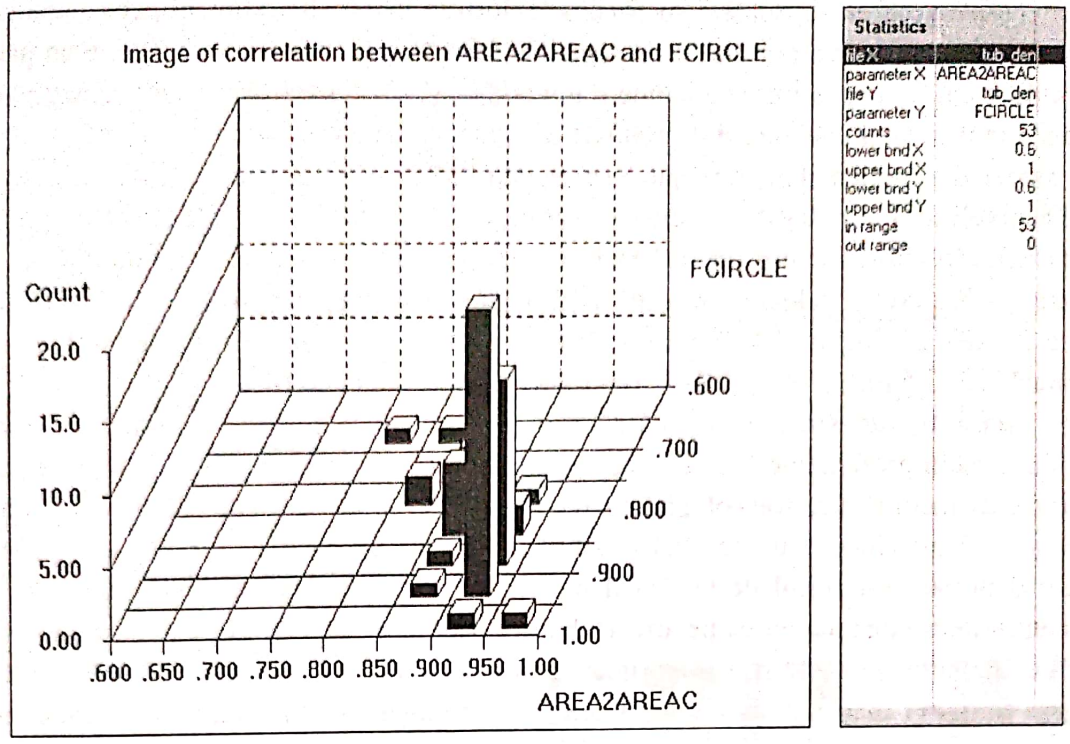


Fig. 4.3.7. Reprezentare tridimensională a distribuției diametrelor echivalente în raport cu factorul de convexitate şi factorul de formă, însoțită de informațiile numerice standard specifice Zeiss KS400.

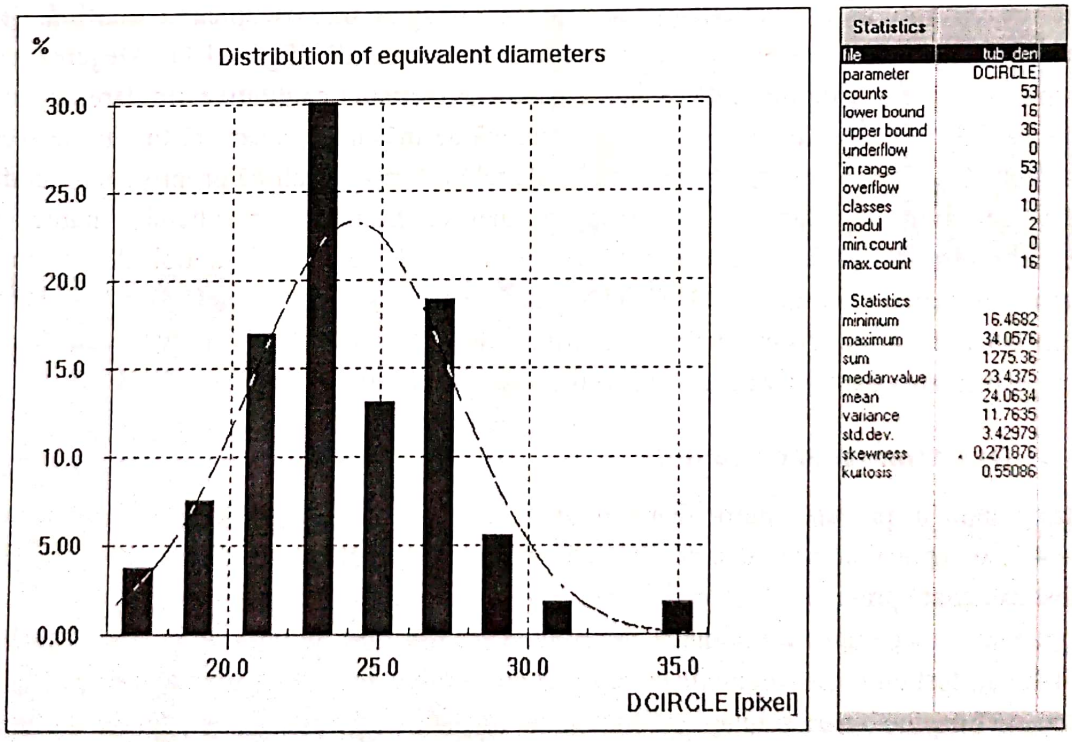


Fig. 4.3.8. Histogramă de frecvențe relative reprezentând, în exprimare procentuală, distribuția diametrelor echivalente, însoțită de informațiile statistice standard specifice Zeiss KS400.

Scanned with OKEN Scanner

Transformarea în unități metrice a valorilor în pixeli obținute pentru diametrele echivalente se realizează pornind de la faptul că imaginea inițială (fig. 4.3.1) este un pătrat cu latura de 4 cm, transpus într-o imagine digitală de 512x512 pixeli. Pe de altă parte, imaginea inițială este rezultatul unei măriri de 1000 de ori a preparatului microscopic, ceea ce înseamnă că unui micron din preparatul microscopic îi corespund 12.8 pixeli (deoarece 512 pixeli/(4 cm /1000) = 12.8 pixeli/μm). Cu aceste informații, conversia diametrelor echivalente din pixeli (a se vedea valorile numerice din fig. 4.3.8) în microni conduce la valoarea minimă de 1.28 μm (pentru 16.46 pixeli), valoarea maximă 2.65 μm (pentru 34.05 pixeli) și, respectiv, valoarea medie de 1.87 μm (pentru 24.06 pixeli), datele obținute încadrându-se în domeniul raportat în literatură (0.9-2.5 μm) (Avery, 1994, Ten Cate, 1998, Căruntu, 2001 a).

Acest tip de structuri tubulare, prin caracteristicile histoarhitectonice – respectiv continuitatea în profunzime (perpendicular pe secțiune), satisfac cu suficient de bună precizie ipotezele de lucru pentru **stereologie**. Astfel, s-a putut calcula densitatea de volum a tubulilor dentinari, obținându-se valoarea 9.3%, adică într-un volum de 1 mm^3 9.3% este ocupat de tubulii dentinari, iar restul de 90.7% îl reprezintă dentina peritubulară și intertubulară. De asemenea, densitatea numerică pe arie a tubulilor dentinari este de 33125 tubuli/mm^2, care, ca valoare, rămâne identică și în exprimarea densității pe volum, adică 33125 tubuli/mm^3, datorită lungimii mari a tubulilor dentinari. (Dimensiunea tubulilor face ca, în două secțiuni virtuale situate la distanță de 1 mm, să se identifice aproximativ aceeaşitubuli.)

4. 4. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU STRUCTURI TUBULARE ÎN SECȚIUNE LONGITUDINALĂ. ILUSTRARE LA NIVELUL DENTINEI

Al treilea exemplu a avut ca punct de pornire tot o imagine microscopică de **dentină**, de data aceasta însă **tubulii dentinari** fiind în secțiune longitudinală (fig. 4.4.1). Alegerea acestui exemplu a fost motivată de intenția de a evalua caracteristici cantitative similare cu cele din secțiunea 4.3, utilizând însă o cu totul altă tehnică de măsurare, adecvată tipului de secțiune considerat. Pe lângă prezentarea elementelor de tehnică morfometrică proprii studiului de caz curent, am avut în vedere și posibilitatea unei confruntări a rezultatelor numerice cu informația obținută în cazul exemplului din secțiunea 4.3 a acestui capitol. Acest procedeu a fost implementat prin macroul **DENTIN** (vezi textul sursă din Anexa II.3), care a exploatat posibilitatea de a determina valorile mărimii denumite **diametru interceptat**, introdusă pentru exprimarea cantitativă a dimensiunii transversale a tubulilor (Căruntu, 2002).

4.4.1. DESCRIEREA PROCEDURII

Pentru a intra în posesia tuturor diametrelor interceptate ce pot fi construite pe imaginea din fig. 4.4.1, am aplicat un procedeu de segmentare care a condus la imaginea din fig. 4.4.2. Ulterior, au fost executate prelucrări binare adecvate formelor tubulare. Astfel, pe imaginea din fig. 4.4.2 s-a practicat o operație morfologică de erodare, cu rezultat în fig. 4.4.3. Din imaginea din fig. 4.4.3 au fost eliminate regiunile cu arii nesemnificative, obținându-se imaginea din fig. 4.4.4. Pe această imagine binară s-au aplicat consecutiv operațiile morfologice de umplere a găurilor, și, respectiv, de închidere, care au condus la imaginea din fig. 4.4.5. Acesteia i s-a asociat o bordură de culoare neagră (imaginea din fig. 4.4.6). Respectiva bordură a fost necesară pentru a permite desfășurarea măsurătorilor care, obligatoriu, trebuie să aibă ca punct de plecare o zonă de negru.

Scanned with OKEN Scanner

Astfel, imaginea din fig. 4.4.6 este similară cu cea din fig. 4.4.5 (mai puțin existența bordurii), și are rol de mască pentru măsurare sub Zeiss KS400.

Fig. 4.4.1. Dentină – tubuli dentinari în secțiune longitudinală; imagine inițială pentru macroul DENTIN.

Fig. 4.4.2. Imagine binară rezultată prin operația de segmentare a imaginii din fig. 4.4.1.

Fig. 4.4.3. Imagine binară reprezentând rezultatul unor operații succesive de erodare verticală.

Scanned with OKEN Scanner

Fig. 4.4.4. Imagine binară reprezentând rezultatul operației de eliminare a ariilor nesemnificative pe imaginea din fig. 4.4.3.

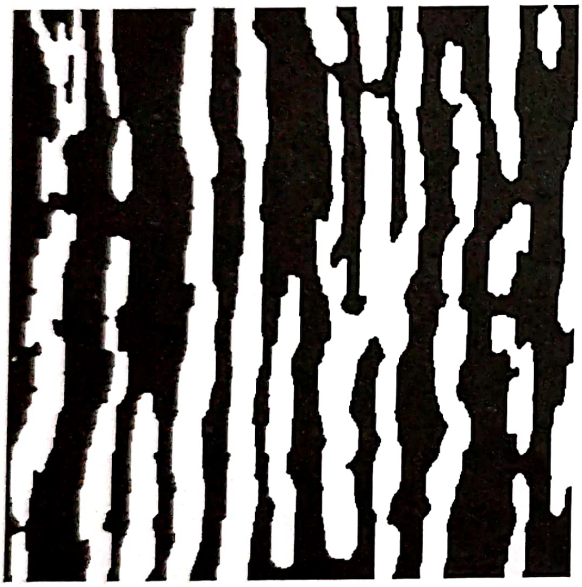


Fig. 4.4.5. Imagine binară rezultată în urma unor operații succesive de umplere a găurilor și închidere aplicate pe imaginea din fig. 4.4.4.

Fig. 4.4.6. Imagine-mască pentru măsurători, similară cu imaginea din fig. 4.4.5, având, în schimb, aplicată o bordură neagră.

Scanned with OKEN Scanner

Dată fiind imaginea-mască creată (cu dimensiuni de 512x512 pixeli), un **diametru interceptat** (variabila *DIAMETER*) a fost definit drept orice coardă orizontală, delimitată de conturul unui tubul, aparţinând oricăreia dintre cele 512 linii ce compun imaginea. Această imagine, abordată morfometric drept unic câmp, a permis aplicarea funcţiei Zeiss KS400 specializată în determinarea numărului total de **puncte de intercepţie** (ce constituie valoarea curentă a variabilei *FLDINTERCEPT*), corespunzând tuturor frontierelor negru-alb, alb-negru, identificate pe fiecare din cele 512 linii orizontale, prin parcurgerea imaginii de la stânga la dreapta şi de sus în jos.

Prin realizarea unei proceduri iterative de erodare orizontală a zonelor albe (echivalente tubulilor), cu câte un singur pixel la fiecare frontieră negru-alb, alb-negru, urmată de aplicarea funcţiei Zeiss KS400 de intercepţie, s-a putut calcula numărul total de diametre interceptate. Determinarea automată a acestora a condus la valori pare cuprinse între 2 pixeli şi 72 pixeli – când procedura iterativă se încheie, deoarece imaginea nu mai conţine nici o zonă albă, adică variabila *FLDINTERCEPT* devine 0. Imaginea din fig. 4.4.7 ilustrează o etapă de pe parcursul desfăşurării procesului iterativ. Iteraţia k ($k \geq 2$) a procedurii furnizează numărul total de diametre interceptate cu dimensiunea $DIAMETER(k) = 2k\text{-}2$ pixeli, existente în fig. 4.4.1, pe baza egalităţii:

$$DIAM_COUNT(k) = \frac{DIFF_FLDINTER(k)}{2}, \; k \geq 2,$$

unde:

$$DIFF_FLDINTER(k) = FLDINTERCEPT(k) - FLDINTERCEPT(k-1), \; k \geq 2.$$

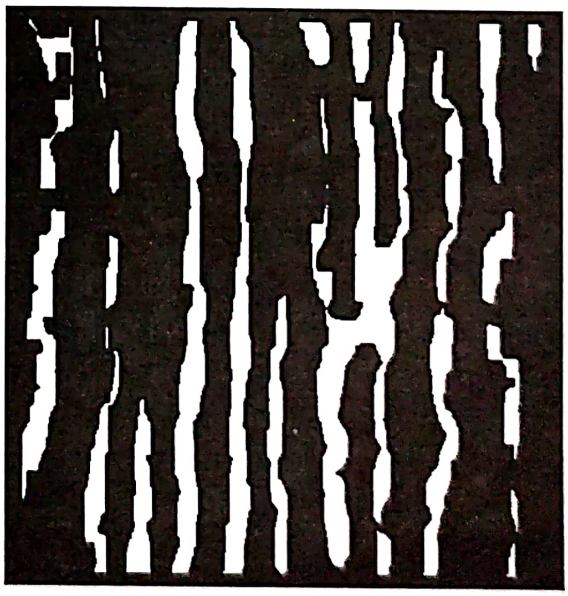


Fig. 4.4.7. Imagine binară rezultată în timpul desfăşurării iteraţiei, reprezentând rezultatul a 5 erodări succesive aplicate pe imaginea din fig. 4.4.6.

Ca urmare a măsurătorilor automate, informaţiile numerice rezultate din evoluţia algoritmului au fost înregistrate într-o bază de date nou creată, de unde au fost extrase şi afişate de Zeiss KS400 într-o fereastră specială sub formă tabelară, din care fig. 4.4.8 reproduce primele 35 de linii (corespunzând primelor 35 de iteraţii). Ulterior, după cum se va vedea mai jos, informaţiile din baza de date au fost de asemenea extrase în scopul realizării unei reprezentări grafice (histogramă).

Scanned with OKEN Scanner

Data List - dentin (4.37)

File Edit Options Help

No	FLDINTERCEPT #	DIFF_FLDINTER #	DIAM_COUNT #	DIAMETER pixel
1	9316.00	0.00	0.00	0.00
2	9286.00	30.00	15.00	2.00
3	9160.00	126.00	63.00	4.00
4	9004.00	156.00	78.00	6.00
5	8662.00	342.00	171.00	8.00
6	8270.00	392.00	196.00	10.00
7	7692.00	578.00	289.00	12.00
8	7208.00	484.00	242.00	14.00
9	6682.00	526.00	263.00	16.00
10	5780.00	902.00	451.00	18.00
11	5102.00	678.00	339.00	20.00
12	4018.00	1084.00	542.00	22.00
13	2718.00	1300.00	650.00	24.00
14	1772.00	946.00	473.00	26.00
15	906.00	866.00	433.00	28.00
16	522.00	384.00	192.00	30.00
17	296.00	226.00	113.00	32.00
18	228.00	68.00	34.00	34.00
19	190.00	38.00	19.00	36.00
20	160.00	30.00	15.00	38.00
21	136.00	24.00	12.00	40.00
22	124.00	12.00	6.00	42.00
23	114.00	10.00	5.00	44.00
24	98.00	16.00	8.00	46.00
25	94.00	4.00	2.00	48.00
26	88.00	6.00	3.00	50.00
27	38.00	50.00	25.00	52.00
28	38.00	0.00	0.00	54.00
29	38.00	0.00	0.00	56.00
30	38.00	0.00	0.00	58.00
31	38.00	0.00	0.00	60.00
32	38.00	0.00	0.00	62.00
33	38.00	0.00	0.00	64.00
34	38.00	0.00	0.00	66.00
35	38.00	0.00	0.00	68.00

Fig. 4.4.8. Tabel cu rezultatele numerice ale măsurătorilor realizate prin macroul DENTIN.

4.4.2. INTERPRETAREA REZULTATELOR MORFOMETRICE

Diametrele interceptate au avut valori situate în domeniul 2-72 pixeli. Informațiile numerice obținute au putut fi caracterizate global prin **media ponderată** a diametrelor interceptate, cu valoarea 21.03 pixeli, calculată sub forma:

$$\left(\sum_{k\geq 2} DIAM_COUNT(k)\times DIAMETER(k)\right)/\left(\sum_{k\geq 2} DIAM_COUNT(k)\right),$$

care, din punct de vedere statistic, are semnificația de medie empirică (speranță matematică) pentru întreaga mulțime de diametre măsurate.

Rezultatele obținute în urma executării macroului **DENTIN** au fost organizate grafic în histograma de frecvențe relative din fig. 4.4.9, care evidențiază corespondența dintre fiecare valoare numerică a diametrului interceptat $DIAMETER(k)$ (privit drept variabilă aleatoare empirică) și exprimarea procentuală a numărului de astfel de diametre:

$$DIAM_COUNT\%(k) = DIAM_COUNT(k)/\left(\sum_{i\geq 2} DIAM_COUNT(i)\right),\ k\geq 2.$$

Se remarcă faptul că diametrele interceptate au avut valori situate și în afara domeniului 16-36 pixeli, identificat pentru diametrele echivalente cu care s-a operat în exemplul din secțiunea 4.3. Această constatare se justifică prin următoarele **caracteristici** ale tubulilor dentinari, în condițiile analizei dimensiunilor acestora, prin prisma unei imagini 2D:

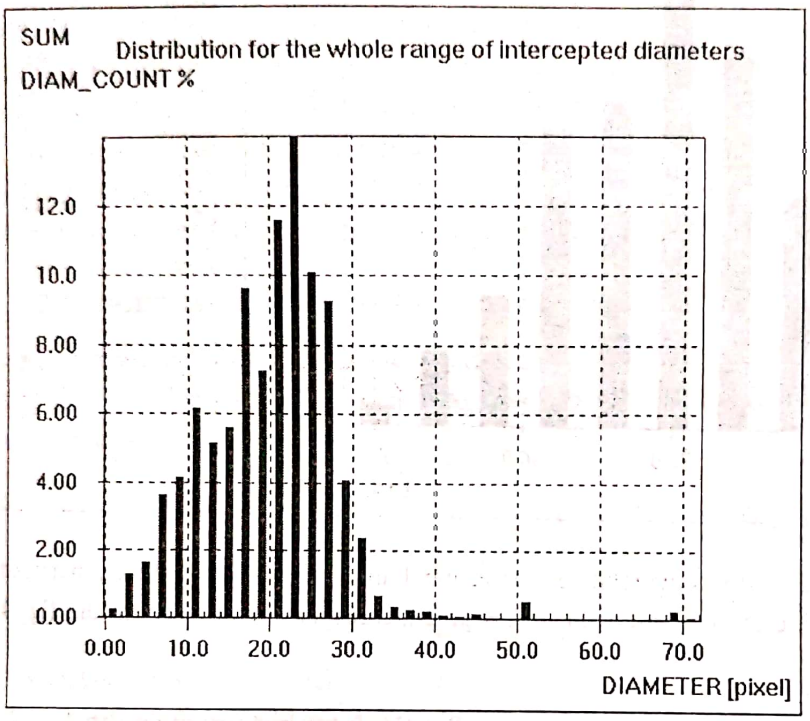
(i) secțiunea longitudinală a unui tubul dentinar poate fi incompletă, datorită includerii doar parțiale a traseului tubulului în planul secțiunii (ceea ce s-a concretizat în existența unor zone albe înguste în fig. 4.4.6 și, respectiv, în găsirea unor valori mici pentru diametrul interceptat);

Scanned with OKEN Scanner

(ii) dimensiunile unui tubul dentinar pot să apară mai mari decât în realitate, datorită existenței unor anastomoze cu tubulii vecini (ceea ce s-a concretizat în existența unor zone albe foarte late în fig. 4.4.6 și, respectiv, în găsirea unor valori mari pentru diametrul interceptat);

(iii) prelucrarea preparatelor, prin tehnica de demineralizare, poate altera dimensiunile reale ale tubulilor dentinari în ambele sensuri (mărire sau reducere artificială), prin estomparea limitelor precise dintre interiorul tubulului și dentina adiacentă.

Operațiile binare utilizate pentru construirea imaginii din fig. 4.4.6 au acționat în sensul reducerii valorilor extreme ale diametrelor interceptate, eficiența lor dovedindu-se cu precădere în raport cu (iii).



Statistics	
file	dentin
parameter	DIAMETER
counts	37
lower bound	0
upper bound	72.1
underflow	0
in range	37
overflow	0
classes	36
modul	2.00278
min.count	0
max.count	650
Statistics	
minimum	0
maximum	72
sum	1332
medianvalue	35.0486
mean	36
variance	468.667
std.dev	21.6487
skewness	0
kurtosis	-1.20175

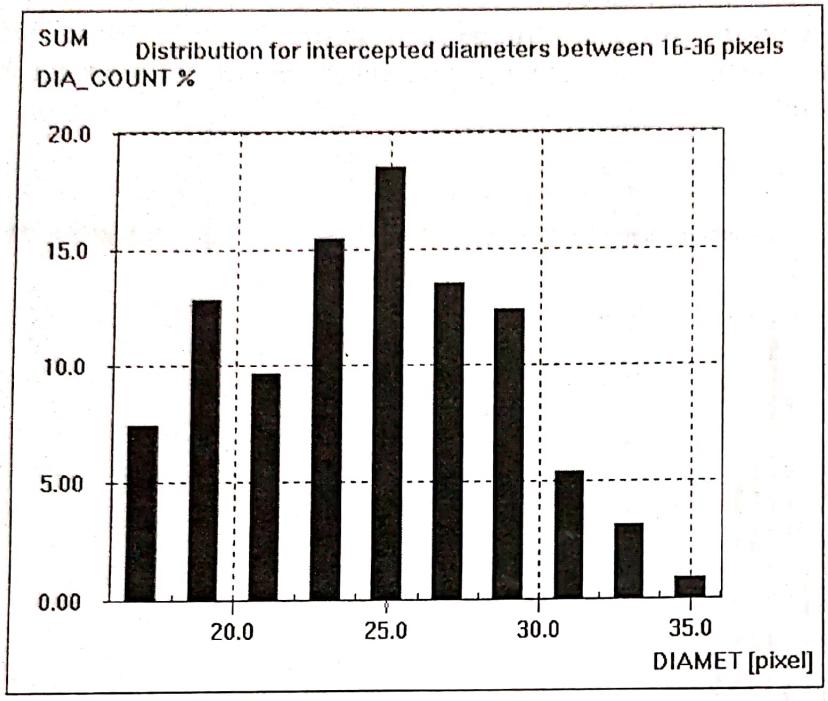
Fig. 4.4.9. Histogramă de frecvențe relative reprezentând, în exprimare procentuală, distribuția tuturor diametrelor interceptate, corespunzătoare imaginii din fig. 4.4.1.

Deoarece macroul **DENTIN** implică un grad înalt de complexitate, am fost interesați de obținerea unei confirmări privind corectitudinea rezultatelor, prin intermediul unei alte tehnici. Astfel, am comparat valoarea de 21.03 pixeli (media ponderată a diametrelor interceptate), dată de macroul **DENTIN,** cu rezultatul de 20.55 pixeli furnizat de funcția Zeiss KS400 *FLDMEANCHORD*, care calculează direct media tuturor diametrelor (corzilor) din fig. 4.4.6. Deși *FLDMEANCHORD* returnează media diametrelor interceptate, utilizarea sa nu poate substitui algoritmul proiectat și implementat prin macroul **DENTIN**, deoarece *FLDMEANCHORD* nu determină numărul de diametre, corespunzând fiecărei valori (pare, din domeniul 2 – 72 pixeli) în parte.

În baza rezultatelor obținute, a devenit evident faptul că o comparație între măsurătorile efectuate în secțiunea curentă și, respectiv, în secțiunea anterioară este relevantă

Scanned with OKEN Scanner

atunci când domeniul diametrelor interceptate se limitează la intervalul 16 – 36 pixeli. În acest context, am dezvoltat un nou macrou **DENT** (vezi textul sursă din Anexa II.3), care exploatează numai acele rezultate furnizate de către **DENTIN** care se încadrează în intervalul sus-menționat și le organizează în forma histogramei de frecvențe relative din fig. 4.4.10, care reprezintă grafic *DIAM_COUNT%(k)* în raport cu *DIAMETER(k)*.



Statistics	
file	den16_36
parameter	DIAMET
counts	10
lower bound	16
upper bound	36
underflow	0
in range	10
overflow	0
classes	10
modul	2
min.count	34
max.count	650
Statistics	
minimum	16
maximum	34
sum	250
medianvalue	26
mean	25
variance	36.6667
std.dev.	6.0553
skewness	0
kurtosis	-1.22424

Fig. 4.4.10. Histogramă de frecvențe relative reprezentând, în exprimare procentuală, distribuția diametrelor interceptate, încadrate în intervalul 16-36 pixeli, corespunzătoare imaginii din fig. 4.4.1.

Se remarcă faptul că există importante similarități între histogramele din fig. 4.4.9 și fig. 4.4.10. Mai mult, în cazul restrângerii domeniului valoric la 16-36 pixeli, în termeni matematici media diametrelor interceptate este 23.37 pixeli, valoare foarte apropiată de 24.06 pixeli, rezultată în studiul de caz din secțiunea 4.3, ca medie a diametrelor echivalente. Astfel, s-a putut demonstra că *măsurătorile efectuate pe același tip de structuri tubulare, dar în secțiuni diferite, conduc la rezultate apropiate, cu diferențe minore în caracterizarea globală.*

Conversia rezultatelor măsurătorilor în unități metrice decurge identic ca în secțiunea 4.3 deoarece imaginea inițială (fig. 4.4.1) este tot un pătrat cu latura de 4 cm, transpus într-o imagine digitală de 512x512 pixeli, fiind de asemenea rezultatul unei măriri de 1000 de ori a preparatului microscopic. Având în vedere faptul că valoarea medie în pixeli a diametrelor, calculată în secțiune transversală și, respectiv, longitudinală, diferă prin mai puțin de 3%, este evident că și exprimarea în microni va conduce la valori practic egale pentru media diametrelor evaluată pentru cele două tipuri de secțiuni.

Spre deosebire de studiul de caz anterior, bazat pe secțiune transversală, utilizarea secțiunii longitudinale nu permite efectuarea unor determinări stereologice credibile, întrucât repetarea virtuală a secțiunii, în profunzime, ar pune în evidență alți tubuli dentinari.

Scanned with OKEN Scanner

4. 5. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU REGIUNI DE INTERES ÎN INTERIORUL UNUI CÂMP. ILUSTRARE PRINTR-UN ADENOCARCINOM ENDOMETRIAL

Acest exemplu se bazează pe o imagine microscopică de **endometru**, la nivelul căruia se observă plaje de **adenocarcinom** nediferențiat, înconjurate de **stromă** abundentă. Localizarea endometrială și tipul histologic al adenocarcinomului reflectă o situație frecventă în activitatea de diagnostic anatomo-patologic, anume stabilirea gradului de dezvoltare și de invazie pentru adenocarcinoame, în raport cu stroma adiacentă, element considerat după unii autori (Colgan et al., 1983, Baak et al., 1984, Baak et al., 1988) ca un posibil **indicator al prognosticului**. Principiile utilizate în analiza cantitativă computerizată pot fi aplicate însă, după cum am menționat anterior, în orice altă localizare, diferențe de metodologie putând fi necesare în raport cu tipurile histologice de adenocarcinom (bine diferențiat, moderat diferențiat, nediferențiat). Acest procedeu a fost implementat prin macroul **ADK** (vezi textul sursă din Anexa II.4), care a exploatat posibilitatea de a determina dimensiunile totale ale **ariilor tumorale** și respectiv, ale **ariilor stromale**; în plus, s-a realizat evaluarea cantitativă a **ariilor nucleare** totale corespunzătoare insulelor tumorale și stromale, precum și calcularea procentului ocupat de nuclei în interiorul acestora (Căruntu, 2003 a).

4.5.1. DESCRIEREA PROCEDURII

Pe imaginea inițială s-a aplicat, anterior determinărilor automate, un procedeu de creștere a contrastului RGB care a condus la imaginea din fig. 4.5.1, considerată drept imagine de pornire în macroul **ADK**.

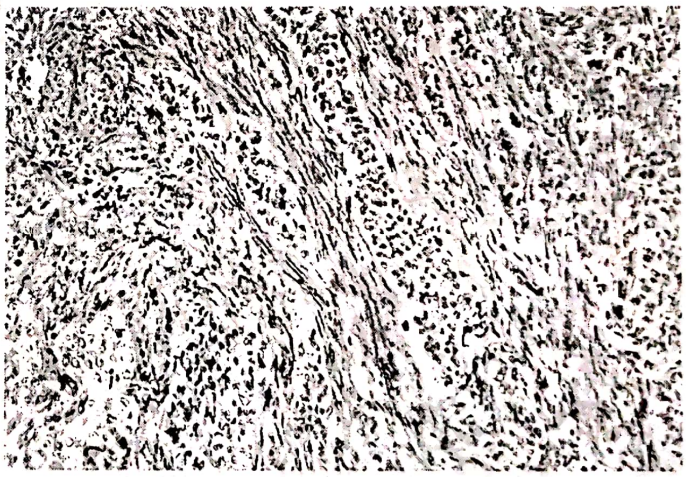


Fig. 4.5.1. Adenocarcinom moderat diferențiat cu stromă abundentă; imagine inițială pentru macroul ADK.

Astfel, în această imagine, elementele morfologice de tipuri diferite capătă culori care, prin transferul în imagine monocromă, se localizează în zone ușor separabile ale histogramei, fapt ce a facilitat segmentarea ulterioară. Pentru delimitarea contururilor celulare tumorale și a țesutului conjunctiv stromal, pe imaginea din fig. 4.5.1 s-a practicat o segmentare de tip prag, în manieră interactivă, cu rezultat în imaginea din fig. 4.5.2. Această imagine a fost supusă unei operații morfologice de umplere a găurilor, rezultând imaginea din fig. 4.5.3. Ulterior, s-a realizat îndepărtarea regiunilor cu arii mai mici de 7000 pixeli^2, obținându-se imaginea din fig. 4.5.4, în care regiunile de interes corespunzătoare zonelor tumorale sunt reprezentate cu alb.

Scanned with OKEN Scanner

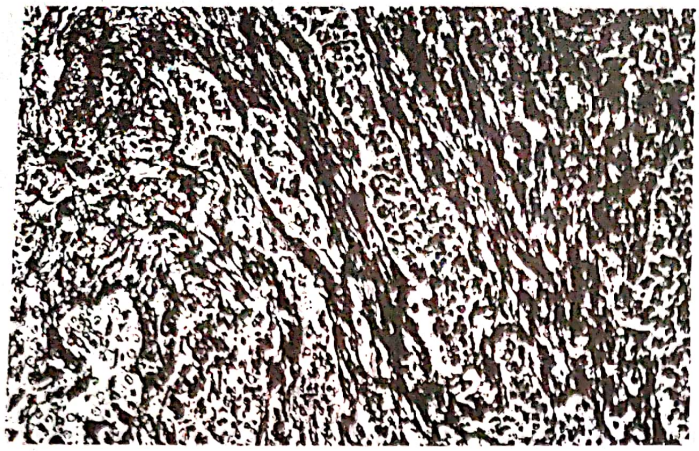


Fig. 4.5.2. Imagine binară rezultată prin operația de segmentare a imaginii din fig. 4.5.1; parametrii selectați interactiv în segmentare au permis marcarea distinctă a insulelor tumorale și a țesutului conjunctiv stromal.

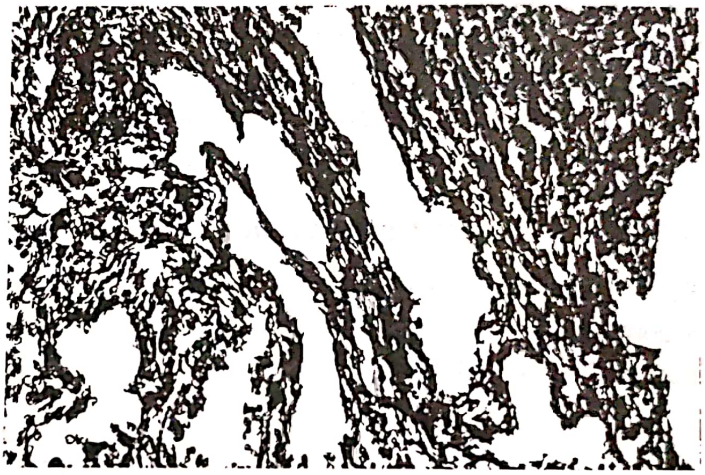


Fig. 4.5.3. Imagine binară reprezentând rezultatul operației de umplere a găurilor pe imaginea din fig. 4.5.2.

Fig. 4.5.4. Imagine binară reprezentând rezultatul operației de eliminare a regiunilor nesemnificative pe imaginea din fig. 4.5.3.

Negativul imaginii din fig. 4.5.4 (obținut prin operația booleană NU) este imaginea binară din fig. 4.5.5, aici (spre deosebire de imaginea din fig. 4.5.4) zonele albe fiind asociate cu stroma adiacentă tumorii.

Scanned with OKEN Scanner

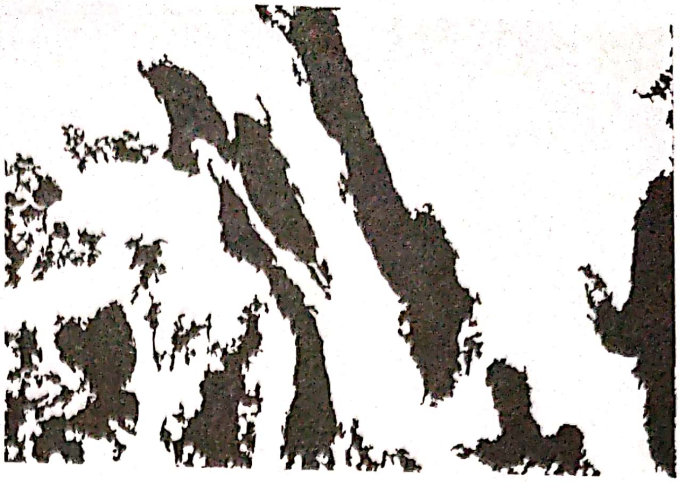


Fig. 4.5.5. Imagine binară reprezentând rezultatul operației booleene NU aplicată pe imaginea din fig. 4.5.4.

Pentru delimitarea tuturor nucleilor (aparținând atât celulelor tumorale, cât și celulelor stromale), pe imaginea din fig. 4.5.1 s-a practicat o a doua segmentare de tip prag, în manieră interactivă (însă cu alți parametri – vezi Anexa II.4), care a condus la imaginea binară din fig. 4.5.6. Întrucât valorile nuanțelor de gri din colorația nucleilor celulelor tumorale și, respectiv, a celulelor stromale nu permit diferențierea acestora, aria nucleară totală pentru zonele tumorale și aria nucleară totală pentru zonele stromale au fost evaluate prin intermediul unor artificii care sunt detaliate după cum urmează.

Pentru zonele tumorale, s-a construit imaginea din fig. 4.5.7, ca diferență dintre mulțimea pixelilor albi din imaginea din fig. 4.5.4 și mulțimea pixelilor albi din imaginea din fig. 4.5.6, astfel încât în fig. 4.5.7 nucleii tumorali apar sub forma unor regiuni de culoare neagră, incluse în regiunile albe (care semnifică ansamblul zonelor tumorale). Drept urmare, aria nucleară tumorală totală s-a calculat scăzând, din aria aferentă pixelilor albi din fig. 4.5.4, aria aferentă pixelilor albi din fig. 4.5.7.

În mod similar, pentru zonele stromale, s-a construit imaginea din fig. 4.5.8, ca diferență dintre mulțimea pixelilor albi din imaginea din fig. 4.5.5 și mulțimea pixelilor albi din imaginea din fig. 4.5.6, astfel încât în fig. 4.5.8 nucleii stromali apar sub forma unor regiuni de culoare neagră, incluse în regiunile albe (care semnifică ansamblul zonelor stromale). Drept urmare, aria nucleară stromală totală s-a calculat scăzând, din aria aferentă pixelilor albi din fig. 4.5.5, aria aferentă pixelilor albi din fig. 4.5.8.

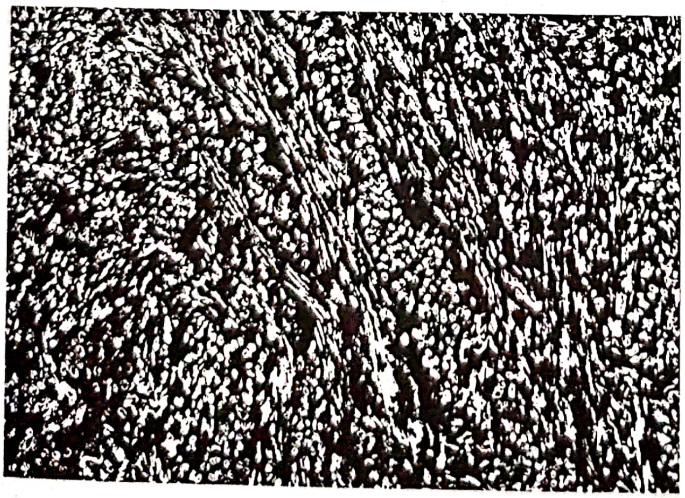


Fig. 4.5.6. Imagine binară rezultată prin operația de segmentare a imaginii din fig. 4.5.1; parametrii selectați interactiv în segmentare au permis marcarea distinctă a nucleilor celulelor tumorale și a nucleilor din stroma adiacentă.

Scanned with OKEN Scanner

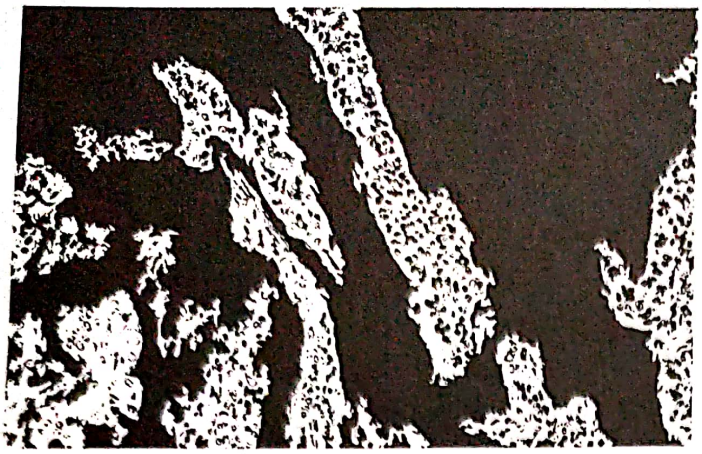


Fig. 4.5.7. Imagine binară rezultată prin operația aritmetică de scădere a imaginii din fig. 4.5.6 din imaginea din fig. 4.5.4.

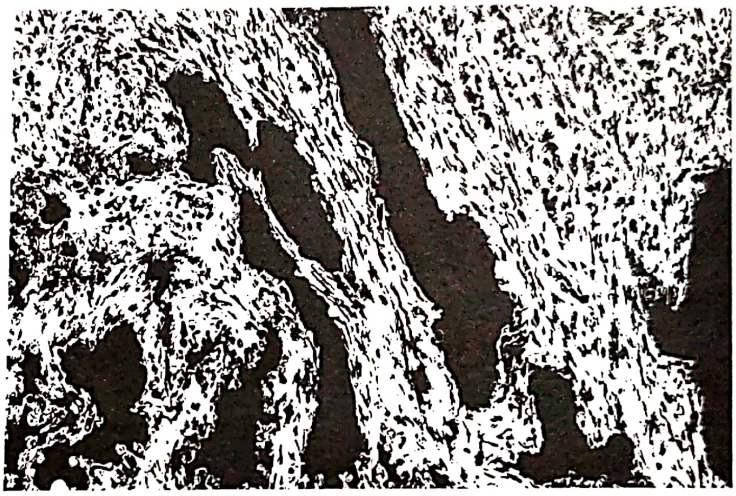


Fig. 4.5.8. Imagine binară rezultată prin operația aritmetică de scădere a imaginii din fig. 4.5.6 din imaginea din fig. 4.5.5.

Pentru măsurători a fost utilizată trăsătura de câmp *FLDAREA* care, conform principiilor de calcul expuse anterior, a fost aplicată succesiv pentru a determina **ariile** zonelor de pixeli albi din imaginile binare (cu rol de mască de măsurătoare) create în figurile 4.5.4, 4.5.6, 4.5.5, 4.5.7.

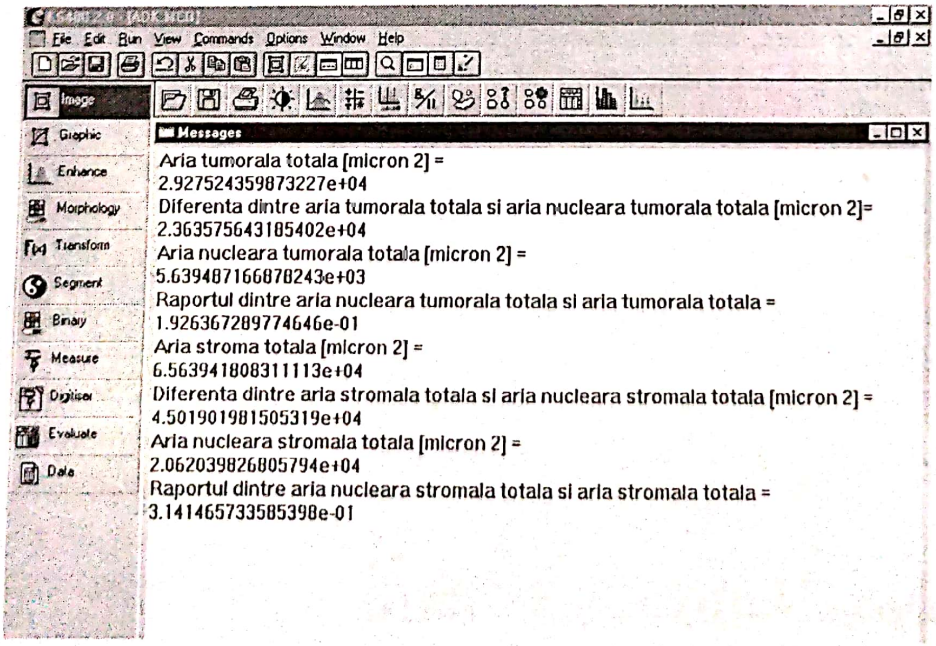


Fig. 4.5.9. Fereastră de mesaje afișând informațiile numerice rezultate din măsurători.

Scanned with OKEN Scanner

Ca urmare a măsurătorilor automate, informațiile numerice rezultate au fost memorate într-o bază de date nou creată, de unde au fost extrase și afișate de Zeiss KS400 într-o fereastră specială de tip mesaj, prezentată în fig. 4.5.9.

4.5.2. INTERPRETAREA REZULTATELOR MORFOMETRICE

Conform informațiilor numerice obținute prin măsurători, se constată că, în câmpul pe care s-a lucrat, aria tumorală totală este de 29275 μm^2 (= 0.029275 mm^2) și aria stromală totală este de 65639 μm^2 (= 0.065639 mm^2), reprezentând 31%, și, respectiv, 69%.

În interiorul ariei tumorale totale, citoplasma celulară ocupă 23635 μm^2 (= 0.023635 mm^2) – adică 80.8%, iar nucleii celulelor reprezintă 5639 μm^2 (= 0.005639 mm^2) – adică 19.2%. De asemenea, în interiorul ariei stromale totale, fibrele de colagen și citoplasma celulelor conjunctive ocupă 45019 μm^2 (= 0.045019 mm^2) – adică 68.6%, iar nucleii celulelor reprezintă 20620 μm^2 (= 0.002062 mm^2) – adică 31.4%.

Considerând că sunt îndeplinite condițiile necesare unor **determinări stereologice**, în sensul că procentele de arie stabilite anterior se regăsesc în secțiuni seriate (datorită structurii repetitive tridimensionale a tumorii), rezultă că, în orice volum din țesutul prelevat, 31% este ocupat de tumoră iar 69% este ocupat de stromă. Pentru zonele tumorale, densitatea de volum a nucleilor este 19.2%, iar densitatea de volum a citoplasmei celulare este 80.8%. Pentru teritoriul stromal, densitatea de volum a fibrelor de colagen și citoplasmei celulelor conjunctive este 68.6%, iar densitatea de volum a nucleilor celulelor stromale este 31.4%.

Faptul că în volumul tumoral total citoplasma celulară reprezintă aproximativ 80%, iar nucleii celulelor tumorale aproximativ 20%, poate fi interpretat ca un **indicator al gradului de diferențiere tumorală**. Tumora este formată din celule mari, coezive între ele, cu o masă citoplasmatică bine reprezentată, dar fără a exprima componente citoplasmatice specifice și fără a avea o delimitare precisă. Putem considera că exprimările procentuale de mai sus reflectă caracteristicile unui adenocarcinom foarte puțin diferențiat, cu o importantă densitate celulară. În acest context, corelând rezultatele obținute cu gradul scăzut de diferențiere, putem privi valorile procentuale sus-menționate ca argumente (morfometrice) de prognostic defavorabil.

Totodată, analizând rezultatele obținute pentru componenta stromală, valoarea procentuală a fibrelor de colagen și citoplasmei celulelor conjunctive, pe de o parte, și valoarea procentuală a nucleilor celulelor conjunctive, pe de altă parte, pot fi interpretate ca o reactivitate tisulară bună. Stroma adiacentă tumorii, bogat celularizată (densitatea de volum a nucleilor reprezentând mai mult de 30%), poate fi astfel considerată un **factor de ameliorare a prognosticului**, spre deosebire de o stromă mult mai colagenizată, cu o prezență redusă a celulelor conjunctive. În evaluarea stromei, un aspect care trebuie subliniat și luat în considerație, tot ca factor de ameliorare a prognosticului, este absența angiogenezei în focarul tumoral, știut fiind că un proces exuberant de angiogeneză facilitează dezvoltarea tumorală.

Raportul dintre volumul tumoral total și volumul stromal total (0.45) poate fi interpretat similar, ca **factor de ameliorare a prognosticului**.

Precizăm faptul că, în formularea unui prognostic pe baza criteriilor morfologice, parametrii de evaluare au o valoare predictivă absolută când sunt luați în considerare individual (componenta tumorală, separat de componenta stromală). În cazul în care se asociază doi parametri, valoarea predictivă devine relativă, gravitatea datorată unui parametru putând fi atenuată prin caracteristicile celui de al doilea parametru. De aceea, este oportună o evaluare pe baza a cât mai multor parametri, riscul de eroare fiind astfel diminuat.

Scanned with OKEN Scanner

În final, subliniem faptul că, în contextul lucrării, scopul exemplului este *descrierea unei modalităţi de investigare morfometrică automată în patologia tumorală*, fără a fi preocupaţi de rezultatele aplicării metodei pe mai multe câmpuri. Din acest motiv, valorile numerice obţinute prin măsurare pe câmpul luat în discuţie au doar o *semnificaţie orientativă*. Pentru obţinerea unor valori numerice relevante în sensul caracterizării morfometrice a problemei de patologie luată în discuţie (inclusiv formularea unor factori de prognostic în manieră procentuală), este necesară selectarea şi examinarea unui eşantion reprezentativ de câmpuri, în conformitate cu principiile statistice.

4. 6. ANALIZĂ MORFOLOGICĂ CANTITATIVĂ PENTRU ELEMENTE DE INTERES DIN INTERIORUL UNEI REGIUNI. ILUSTRARE LA NIVELUL CORPUSCULULUI RENAL

Exemplul propus se bazează pe o imagine de corticală obţinută prin puncţie biopsie renală, prezentând un **corpuscul renal** înconjurat de tubi contorţi. La nivelul acestuia sunt evidente modificări morfologice specifice **glomerulonefritei rapid progresive**, observându-se dezvoltarea unei semilune predominant fibrilare de tip circumferenţial, creşterea importantă a matricei mezangiale şi reducerea spaţiului Bowmann (Căruntu et al., 1997).

Pentru efectuarea măsurătorilor automate, descrise în detaliu în cele ce urmează, am proiectat şi implementat macroul denumit **COREN** (vezi textul sursă din Anexa II.5), scopul urmărit fiind determinarea **ariilor corpusculului renal** şi ale principalelor elemente de interes existente în interiorul acestuia: **capilare**, **nuclei**, **matrice**, **spaţiu Bowmann**. Consecutiv, au fost calculate procentele, din aria totală a corpusculului renal, ocupate de respectivele elemente de interes. Metodologia de realizare a acestui procedeu de analiză cantitativă computerizată poate fi extrapolată şi pentru alte situaţii de morfometrie histologică şi anatomo-patologică, în cazul în care sunt vizate anumite elemente de interes (structuri morfologice) conţinute în interiorul unei regiuni.

4.6.1. DESCRIEREA PROCEDURII

Pe imaginea iniţială s-a aplicat o tehnică de îmbunătăţire a contrastului, rezultând imaginea din fig. 4.6.1. Marcarea interactivă a corpusculului renal a fost realizată cu ajutorul editorului grafic. Conturul desenat în planul grafic a fost suprapus peste imaginea color iniţială, obţinându-se imaginea din fig. 4.6.2. Totodată a fost creată o imagine binară complet neagră, cu rol de suport pentru aplicarea planului grafic (fig. 4.6.3). După aplicarea conturului desenat interactiv, pe imaginea din fig. 4.6.3 a fost efectuată operaţia de umplere a găurilor, marcându-se astfel cu alb întreaga arie a corpusculului renal, inclusiv conturul. Deoarece principiul de măsurare a ariei asociază o valoare nenulă mulţimii de pixeli ce alcătuiesc conturul, şi ţinând cont că măsurătorile au vizat regiunea interioară a acestuia, au fost excluşi pixelii conturului prin operaţia booleeană SAU-EXCLUSIV, rezultând imaginea din fig. 4.6.4.

În scopul delimitării diferenţiate a elementelor de interes, a fost utilizată o procedură de segmentare RGB (*eng.* red – roşu, *eng.* green – verde, *eng.* blue – albastru, vezi Anexa I). Acest tip de segmentare a fost anume ales, deoarece segmentarea pe imaginea monocromă rezultată din fig. 4.6.1 nu asigură acurateţea necesară individualizării structurilor diferite. Parametrii selectaţi în segmentare au fost aleşi, pentru fiecare tip de element, prin aprecierea valorii predominante (în domeniul [0, 255]) a nuanţelor de roşu, verde şi albastru. Astfel, în

Scanned with OKEN Scanner

urma segmentării pentru selectarea zonelor de culoare albă a rezultat imaginea din fig. 4.6.5, în urma segmentării pentru selectarea zonelor de culoare maro a rezultat imaginea din fig. 4.6.6, iar în urma segmentării pentru selectarea zonelor de culoare verde a rezultat imaginea din fig. 4.6.7. O segmentare separată de tip creştere de regiuni a fost aplicată tot pe imaginea din fig. 4.6.2, de data aceasta pentru individualizarea spațiului Bowmann, obținându-se imaginea din fig. 4.6.8.

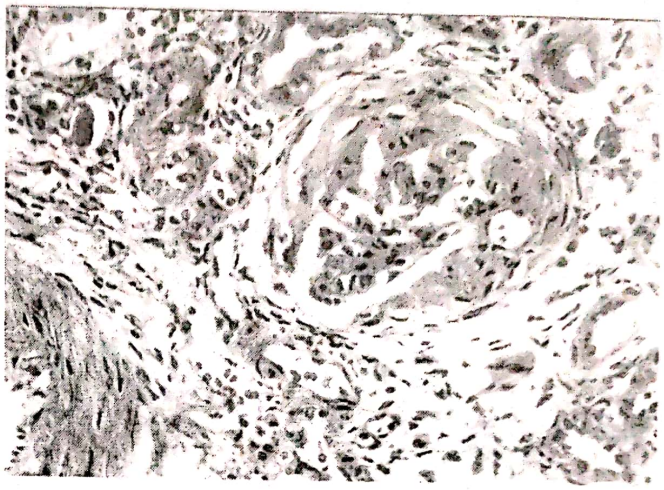


Fig. 4.6.1. Corpuscul renal cu modificări morfologice în cadrul unei glomerulonefrite; imagine inițială pentru macroul COREN.

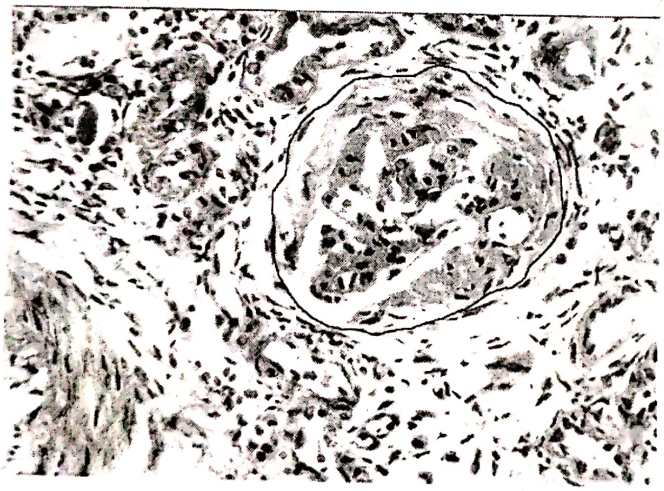


Fig. 4.6.2. Imaginea inițială peste care s-a aplicat conturul corpusculului renal.

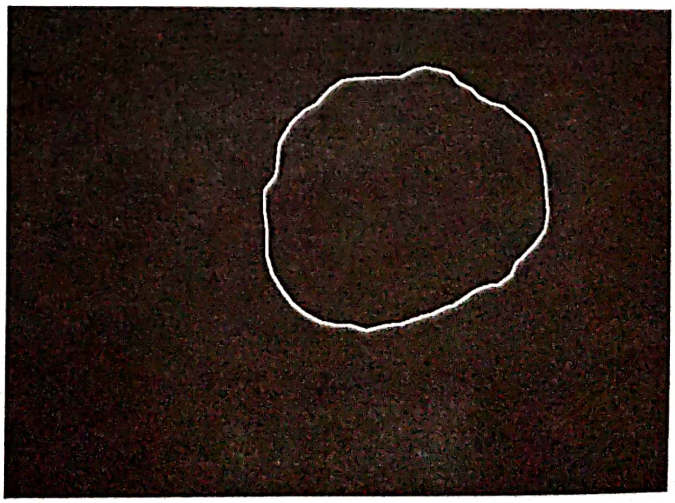


Fig. 4.6.3. Aplicarea conturului corpusculului renal, desenat în planul grafic, peste o imagine binară special creată.

Scanned with OKEN Scanner

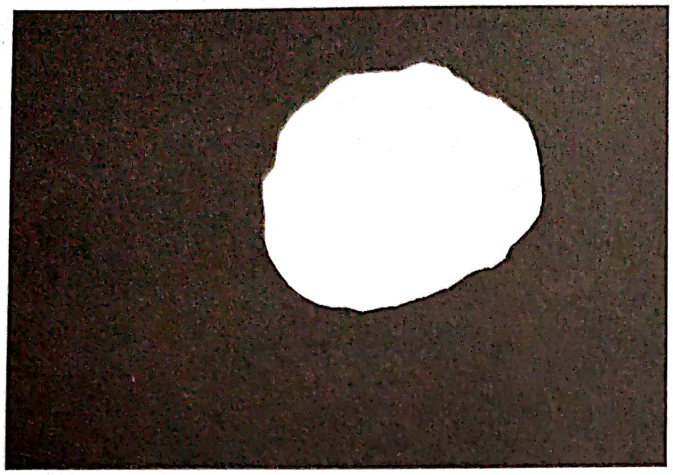


Fig. 4.6.4. Rezultatul operației de umplere a găurilor aplicată pe imaginea din fig. 4.6.3.

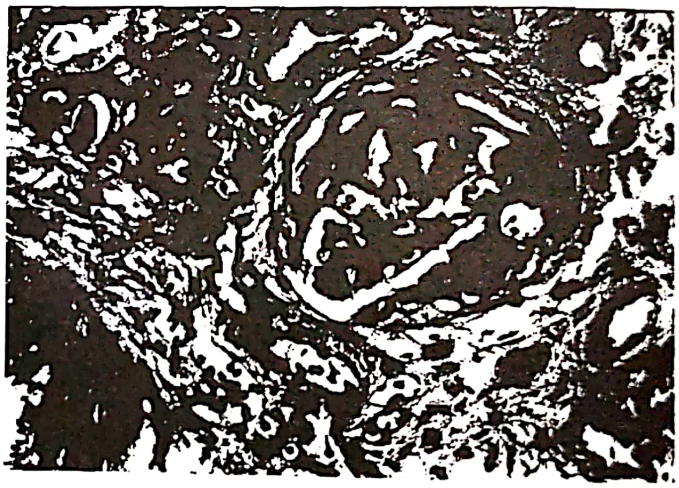


Fig. 4.6.5. Rezultatul operației de segmentare RGB, practicată pentru selectarea zonelor de culoare albă din imaginea inițială.

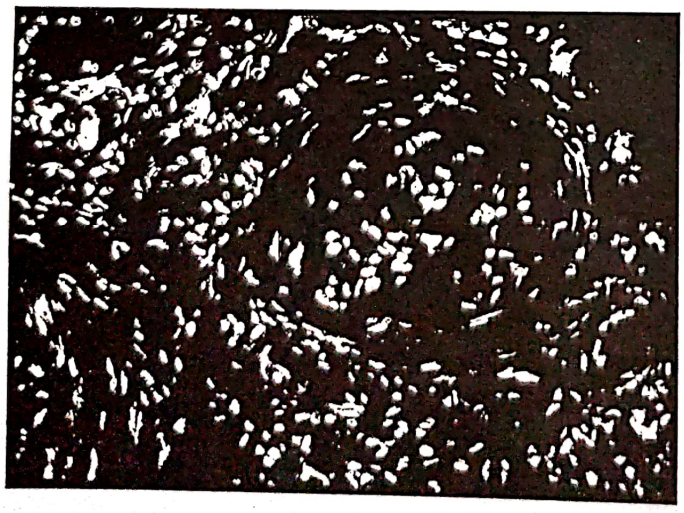


Fig. 4.6.6. Rezultatul operației de segmentare RGB, practicată pentru selectarea zonelor de culoare maro din imaginea inițială.

Scanned with OKEN Scanner

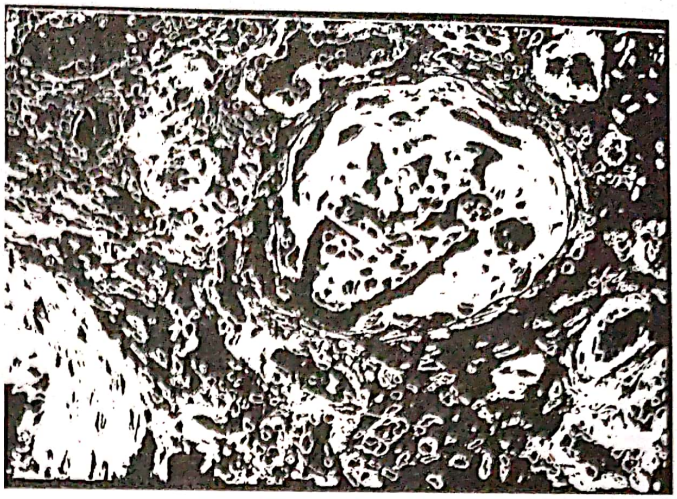


Fig. 4.6.7. Rezultatul operației de segmentare RGB, practicată pentru selectarea zonelor de culoare verde din imaginea inițială.

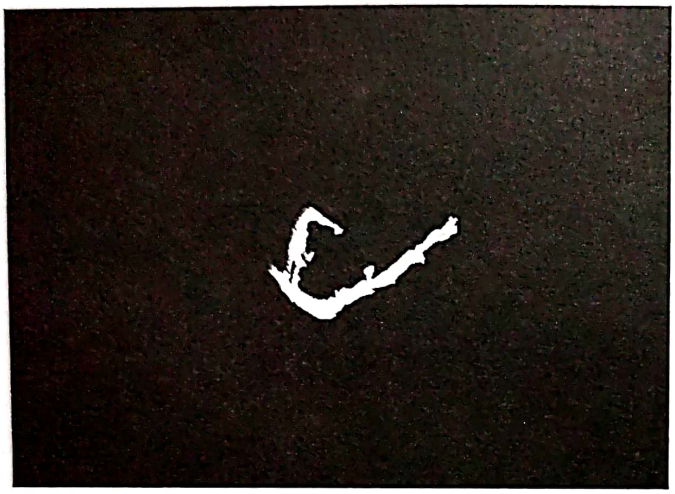


Fig. 4.6.8. Rezultatul operației de segmentare tip creştere de regiuni, practicată pentru selectarea spațiului Bowmann.

Menționăm faptul că operația de segmentare a fost aplicată pe întreaga imagine din fig. 4.6.2, de aceea rezultatul segmentării a inclus şi alte elemente în afara corpusculului renal. Din acest motiv, ulterior segmentării, au fost realizate 4 operații booleene de tip ŞI, în scopul de a selecta, pe rând, acele mulțimi de pixeli care sunt comuni, pe de o parte, imaginii din fig. 4.6.4 (corpusculul renal) şi, pe de altă parte, fiecăreia din cele 4 imagini obținute în urma segmentărilor. Astfel, prin selectarea mulțimii de pixeli albi comuni pentru imaginile din fig. 4.6.4 şi fig. 4.6.5 a rezultat imaginea binară din fig. 4.6.9, în care zonele albe corespund capilarelor reunite (în sens de mulțime) cu spațiul Bowmann. Prin selectarea mulțimii de pixeli albi comuni pentru imaginile din fig. 4.6.4 şi fig. 4.6.8 a rezultat o altă imagine binară, în care zona albă corespunde numai spațiului Bowmann. (Întrucât spațiul Bowmann se află strict în interiorul delimitat prin contur, această imagine coincide cu imaginea din fig. 4.6.8, motiv pentru care nu o mai reproducem.). Imaginea din fig. 4.6.10 corespunde numai capilarelor şi a fost construită ca diferență dintre mulțimea pixelilor albi din imaginea din fig. 4.6.9 şi mulțimea pixelilor albi din imaginea din fig. 4.6.8.

Prin selectarea mulțimii de pixeli albi comuni pentru imaginile din fig. 4.6.4 şi fig. 4.6.6 a rezultat imaginea binară din fig. 4.6.11, în care zonele albe corespund numai nucleilor celulelor mezangiale şi podocitare.

Prin selectarea mulțimilor de pixeli albi comuni pentru imaginile din fig. 4.6.4 şi fig. 4.6.7 a rezultat imaginea binară din fig. 4.6.12, în care zonele albe corespund numai matricei existente în corpuscul, atât la nivel mezangial, cât şi la nivel de crescent.

Scanned with OKEN Scanner

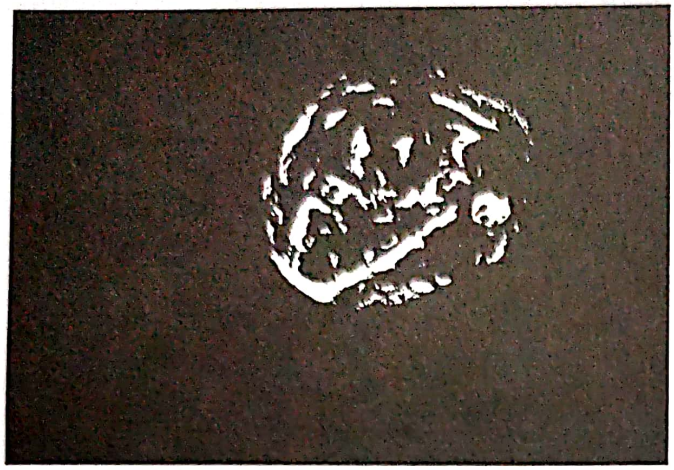


Fig. 4.6.9. Rezultatul aplicării operației booleene ȘI pe imaginile din fig. 4.6.4 și fig. 4.6.5.

Fig. 4.6.10. Rezultatul aplicării operației aritmetice de scădere pe imaginile din fig. 4.6.9 și fig. 4.6.8.

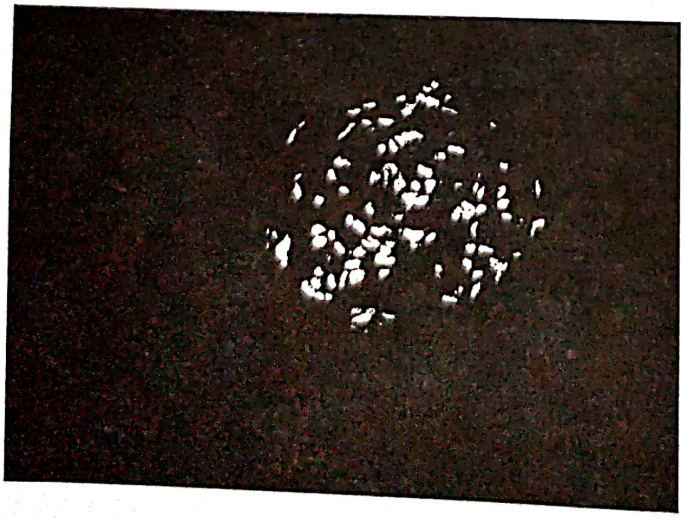


Fig. 4.6.11. Rezultatul aplicării operației booleene ȘI pe imaginile din fig. 4.6.4 și 4.6.6.

Scanned with OKEN Scanner

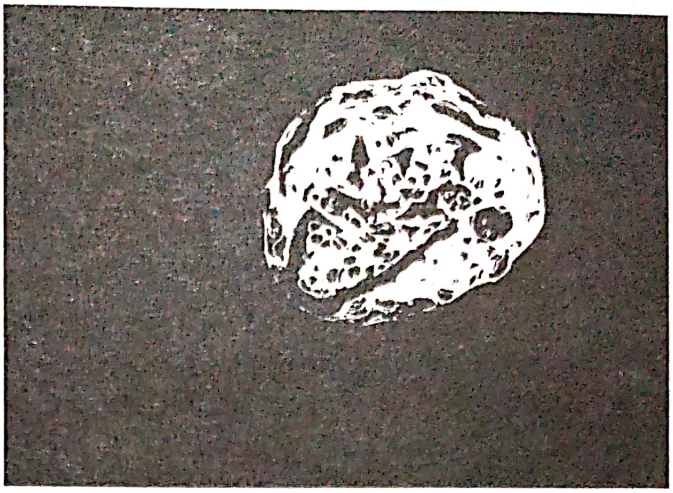


Fig. 4.6.12. Rezultatul aplicării operației booleene ȘI pe imaginile din fig. 4.6.4 și fig. 4.6.7.

Pentru măsurători a fost utilizată trăsătura de câmp *FLDAREA* care, ca și în secțiunea 4.5, a fost aplicată succesiv pentru a determina ariile zonelor de pixeli albi din imaginile binare din figurile 4.6.4, 4.6.8, 4.6.10, 4.6.11, 4.6.12, cu rol de imagini-mască în extragerea trăsăturii.

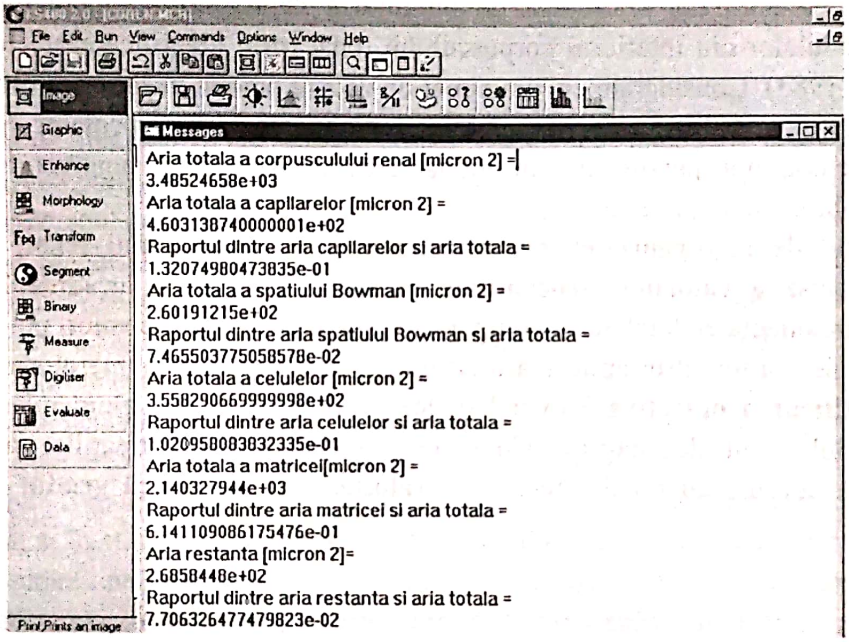


Fig. 4.6.13. Fereastră de mesaje afișând informațiile numerice rezultate din măsurători.

Informațiile numerice obținute ca urmare a determinării automate a ariilor totale ale elementelor de interes au fost memorate (conform modului de operare specific Zeiss KS400) într-o bază de date special creată, de unde au fost preluate în vederea afișării și efectuării altor calcule. Rezultatele de interes au fost scrise în fereastră de mesaje (prezentată în fig. 4.6.13), care asigură comunicația cu utilizatorul.

4.6.2. INTERPRETAREA REZULTATELOR MORFOMETRICE

Conform informațiilor numerice obținute prin măsurători, se constată că aria totală a corpusculului renal este de 3485 μm^2 (= 0.003485 mm^2). În interiorul acesteia, capilarele

Scanned with OKEN Scanner

ocupă o arie totală de 460 μm^2 (= 0.000460 mm^2), celulele ocupă o arie totală de 355 μm^2 (= 0.000355 mm^2), țesutul conjunctiv al matricei mezangiale și al crescentului are o arie totală de 2140 μm^2 (= 0.002140 mm^2), iar spațiul Bowmann – o arie totală de 260 μm^2 (= 0.000260 mm^2). Deoarece suma tuturor ariilor totale ale acestor elemente de interes este de 3216 μm^2 (= 0.003216 mm^2), valoare mai mică decât aria totală a corpusculului, s-a calculat aria restantă, care este de 268 μm^2 (= 0.000268 mm^2). Existența acestei arii restante este datorată unor elemente care nu au fost puse în evidență prin segmentare și nu au făcut obiectul măsurătorilor – respectiv hematiile din interiorul capilarelor, precum și zonele de contur care delimitează capilarele și matricea.

Din punct de vedere procentual, capilarele reprezintă 13.2% din aria totală a corpusculului, celulele – 10.2%, matricea – 61.4%, spațiul Bowman – 7.4%, iar elementele reziduale – 7.7%.

Puncția biopsie renală permite stabilirea tipului de afectare glomerulară, conform criteriilor utilizate pe plan european și mondial. Literatura de specialitate citează abordări morfometrice ale corpusculului renal, rezultatele obținute prin stereologie fiind axate pe densitatea de volum a matricei mezangiale, densitatea de suprafață a membranei bazale a capilarelor glomerulare, densitatea de lungime a capilarelor glomerulare și densitatea numerică a celulelor din interiorul corpusculului renal (prin identificarea nucleilor acestora) (Collan et al, 1983). Considerăm că, din punct de vedere practic, rezolvarea prin morfometrie automată a determinărilor dimensiunilor corpusculului renal și elementelor componente ale acestuia deschide noi perspective în modalitatea de analiză a puncției biopsie renală, în vederea stabilirii unui prognostic.

Pornind de la organizarea histoarhitectonică a corpusculului renal, este necesară evaluarea precisă a **valorilor procentuale normale** în care se încadrează componenta epitelială, componenta endotelială și componenta mezangială. Aceste valori pot fi considerate ca având rol de martor. Prin aplicarea morfometriei automate în glomerulonefrite, se poate realiza **cuantificarea obiectivă** a extinderii leziunilor proliferative glomerulare, fie că este vorba de celule, sau de matrice. Consecutiv, compararea informațiilor numerice care caracterizează corpusculii renali afectați cu valorile normale indică **gradul de severitate** a leziunilor.

Scanned with OKEN Scanner

Scanned with OKEN Scanner

ANEXA I

PARTICULARITĂŢI ÎN PROCESAREA IMAGINILOR COLOR

În procesarea de imagini şi analiza cantitativă asistate de calculator dedicate morfologiei microscopice, imaginile care fac obiectul investigaţiei sunt **imagini color**, iar achiziţionarea şi afişarea lor pe ecranul calculatorului se fac, de asemenea, sub formă color. Diferitele programe de prelucrare a imaginilor au capacitatea de a lucra deopotrivă pe imagini color sau pe imagini în nuanţe de gri, acestea din urmă rezultând ca urmare a conversiei celor dintâi. Manipularea imaginilor color are, însă, avantajul de a oferi observatorului uman mai multe informaţii vizuale, importante în morfometria computerizată pentru a discerne, într-o imagine, regiunile de interes.

Principiile teoretice şi aplicaţiile practice prezentate pe parcursul acestei lucrări s-au axat pe **imagini monocrome**, deoarece costurile de tipărire a acestui volum nu au permis includerea imaginilor color. Trebuie însă subliniat faptul că imaginile iniţiale, utilizate atât în exemplele din capitolul 2, cât şi în studiile demonstrative din capitolul 4, au fost imagini color care, ca primă etapă în prelucrare, au fost transformate în imagini cu nuanţe de gri. De aceea, considerăm oportună prezentarea câtorva elemente esenţiale privind **teoria culorilor** şi *implicaţiile acesteia în metodologia procesării asistate de calculator a imaginilor*. Pentru o documentare suplimentară în scopul aprofundării acestui subiect, pot fi consultate următoarele lucrări de referinţă (Jain, 1989, Russ, 1998, Alvira, 2000, Shapiro, Stockman, 2001, Gonzales, Woods, 2002).

• SISTEME DE CODIFICARE A CULORILOR

Există numeroase moduri de a descrie culorile. Când un fascicul de lumină albă trece printr-o prismă de sticlă, ochiul uman percepe un spectru color format din roşu, portocaliu, galben, verde, albastru, indigo şi violet. Culorile roşu, albastru şi verde sunt **culori primare**, iar un amestec al acestora realizează culoarea albă. Celelalte culori provin din combinarea a doi sau mai mulţi pigmenţi (culori). De exemplu, sunt considerate **culori secundare** roşul magenta (combinaţie de roşu şi albastru), albastrul cyan (combinaţie de albastru şi verde) şi galbenul (combinaţie de roşu şi verde).

Pentru ochiul uman, senzația de culoare este rezultatul perceperii unei radiații electromagnetice cu lungime de undă cuprinsă aproximativ între 400 nm și 700 nm (respectiv o frecvență cuprinsă aproximativ între 4 x 10^{14} Hz și între 7.5 x 10^{14} Hz) care definește spectrul vizibil. Conform textelor de specialitate (de exemplu Shapiro, Stockman, 2001), în spectrul vizibil, culorilor roșu, verde și albastru le corespund lungimile de undă 700 nm, 546.1 nm și, respectiv, 435.8 nm. **Lumina albă** cuprinde în mod uniform toate frecvențele (sau lungimile de undă) ale spectrului vizibil. Obiectele apar colorate datorită faptului că acestea absorb radiațiile luminoase dintr-un anumit interval de frecvență, reflectându-le pe celelalte. Domeniul de frecvență al radiațiilor reflectate este cel ce definește culoarea obiectului. De aceea, culoarea unui obiect pe care se proiectează lumina albă poate apărea diferită atunci când asupra acelui obiect se proiectează o lumină colorată.

Imaginile preparatelor microscopice de morfologie normală și patologică, colorate în hematoxilină-eozină, acoperă de obicei următoarele zone din spectrul vizibil: 400-510 nm (nuanțe de albastru), 515-560 nm (nuanțe de verde) și 590-660 nm (nuanțe de roșu). Aceste informații sunt absolut necesare pentru stabilirea punctului de echilibru (în raport cu lumina albă) în funcționarea camerei video pentru achiziționarea imaginilor.

Pentru majoritatea software-urilor și a tipurilor de monitoare utilizate în Europa definirea culorilor se bazează pe cantitatea de roșu, verde și albastru din spectru, sistem cunoscut sub denumirea de RGB (*eng.* red, green, blue). **Sistemul RGB** este un sistem color "aditiv", în sensul că permite crearea culorilor prin adăugarea componentelor la culoarea neagră. Pentru fiecare pixel, dintr-o **imagine color**, cele trei componente **RGB** se caracterizează printr-o intensitate luminoasă, luminozitate sau strălucire. Când, pe un pixel al monitorului, toate cele trei culori prezintă aceeași intensitate, culoarea rezultată este gri, în gama de la negru (cea mai joasă intensitate) la alb (cea mai înaltă intensitate).

Codarea tricromatică RGB specifică sistemelor digitale utilizează 3 octeți (3 x 8 = 24 biți), astfel că fiecărui pixel din imaginea color de 24 de biți (sau 3 octeți) îi sunt asociați câte 8 biți (sau 1 octet) pentru R, G, și, respectiv, B. În baza celor discutate în capitolul 2, subsecțiunea 2.1.3, utilizarea unui octet pentru fiecare din R, G, B, înseamnă existența a 2^8 = 256 de niveluri pentru fiecare culoare primară, ceea ce înseamnă că, în sistemul RGB, pot fi formate 256x256x256 = 16777216 culori. Referirea lor drept culori în sensul larg al cuvântului este inoportună, întrucât ochiul uman nu le poate diferenția pe toate. De aceea utilizarea lor trebuie înțeleasă numai într-un context tehnic de operare cu imagini digitale, unde se folosește termenul de **cod de culoare**.

Cu această convenție de reprezentare, pentru stocarea unei imagini digitale, fiecare pixel din imagine necesită 24 de biți. Ordinea în care apar în memorie cei trei octeți alocați pentru R, G și, respectiv, B nu prezintă importanță din punct de vedere teoretic, ci doar pentru unele aplicații de programare care exploatează direct codul de culoare.

Intensitatea fiecărei componente RGB variază de la 0 la 255 și caracterizarea culorii unui pixel se poate realiza printr-un triplet de forma (intensitate_R, intensitate_G, intensitate_B). Culorilor R, G, B, așa cum sunt ele definite de spectru, le corespund tripletele (255, 0, 0), (0, 255, 0) și, respectiv, (0, 0, 255). Culorile secundare roșu magenta, albastru cyan și galben sunt descrise de tripletele (255, 0, 255), (0, 255, 255) și, respectiv, (255, 255, 0). Tripletul (0, 0, 0) definește culoarea neagră, tripletul (255, 255, 255) definește culoarea albă, iar pentru orice întreg $0<c<255$, tripletul (c, c, c) definește o nuanță de gri. Utilizarea acestor triplete pentru desemnarea culorii este similară cu modul de precizare a poziției unui punct în spațiul tridimensional, motiv pentru care cele trei componente ale tripletului (adică nivelurile de luminozitate RGB) mai sunt denumite **coordonatele culorii**.

Scanned with OKEN Scanner

Altă modalitate de descriere a culorilor, mai intuitivă, este bazată pe trei parametri: nuanță, saturație și valoare / luminozitate. Acest model este referit ca **sistemul HSV** (*eng.* hue = nuanță, *eng.* saturation = saturație, *eng.* value = valoare/luminozitate/intensitate). Conform unor texte tehnice, dedicate vederii computerizate, de dată foarte recentă, sistemul HSV pare să prezinte o serie de avantaje în prelucrarea imaginilor color comparativ cu sistemul RGB. Prin nuanță se înțelege reprezentarea culorii pure din spectrul vizibil, pentru o anume culoare dată, respectiv lungimea de undă dominantă din spectrul culorii respective, percepută de observatorul uman. Saturația reprezintă cantitatea de alb adăugată la culoarea pură. Valoarea, luminozitatea sau intensitatea reprezintă cantitatea de negru adăugată culorii pure. Acest concept este ușor de reprezentat grafic pe un cerc, în care, de-a lungul periferiei cercului, sunt distribuite următoarele culori: roșu (0^0, 360^0), galben (60^0), verde (120^0), cyan (180^0), albastru (240^0) și magenta (300^0). Din punct de vedere al saturației, periferia cercului reprezintă saturația maximă (adică nuanța conține numai culoare pură), iar centrul cercului reprezintă saturația minimă – valoarea 0 (adică nuanța nu mai conține deloc culoare, ci numai alb). Astfel, mergând de la periferia cercului spre interior, nuanța unei culori se modifică, în sensul că saturația sa în culoare pură scade, și se adaugă alb. Din punct de vedere al intensității, grafic vorbind, aceasta poate fi reprezentată printr-o axă liniară perpendiculară pe suprafața cercului. Pentru o anumită nuanță (cu saturație maximă), atribuirea valorii maxime de intensitate înseamnă transformarea respectivei nuanțe în alb, iar atribuirea valorii minime de intensitate înseamnă transformarea respectivei nuanțe în negru. Altfel spus, creșterea în intensitate este echivalentă cu creșterea strălucirii culorii până la alb complet, iar scăderea intensității este reprezentată de întunecarea până la negru complet. Indiferent de nuanță, dacă saturația este nulă, atunci se obține o nuanță de gri (în funcție de valoarea intensității).

Toate exemplele și aplicațiile dezvoltate în această lucrare folosesc drept imagini inițiale imagini color reprezentate în sistemul RGB. Mediul Zeiss KS400 nu procesează imagini codificate HSV, motiv pentru care nu furnizăm mai multe detalii asupra acestui sistem de codare a culorilor. Cititorilor interesați în cunoașterea sistemului HSV precum și a altor sisteme (cum ar fi Lab, CMY), li se recomandă consultarea bibliografiei (Shapiro, Stockman, 2001).

• CONVERSIA IMAGINILOR RGB ÎN IMAGINI CU NUANȚE DE GRI

O modalitate de a realiza **conversia** unei imagini color RGB într-o imagine cu nuanțe de gri (care este cea implementată în mediul Zeiss KS400) o constituie folosirea mediei aritmetice a celor trei coordonate de culoare R, G, B, asociate unui pixel arbitrar, ca valoare a nivelului de gri asociată acelui pixel. Este evident că, în urma unei atare conversii, pixelii cu nivelul de gri c provin din pixelii colorați (r, g, b) ce se caracterizează prin $[(r + g + b) / 3] = c$, unde notația cu paranteze drepte [ ] desemnează partea întreagă a fracției $(r + g + b) / 3$. Rezultă că un același nivel de gri se atribuie la o multitudine de combinații (r, g, b), ceea ce înseamnă, implicit, o reducere drastică a informației de natură vizuală.

• HISTOGRAME RGB

Histograma unei imagini color de 24 biți se bazează pe definiția din capitolul 2, subsecțiunea 2.2.2, care se aplică individual fiecărei componente R, G sau B, rezultând **trei grafice** distincte. Un grafic se asociază funcției nivel de culoare și exprimă numărul de pixeli din imagine, corespunzător fiecărei luminozități, având pe ordonată numărul de pixeli și pe abscisă cele 256 niveluri de culoare (de la 0 la 255). Altfel spus, se construiește câte

Scanned with OKEN Scanner

o histogramă pentru fiecare coordonată de culoare, sau pentru fiecare din canalele (*eng.* channel) R, G, B. Particularizând pentru fiecare culoare primară, valoarea 0 înseamnă că pixelul respectiv nu conține acea culoare, iar valoarea 255 înseamnă că pixelul respectiv conține acea culoare în cantitate maximă. Așa cum am menționat anterior, în cazul în care, pentru un pixel, toate cele trei culori primare au valoarea 0, pixelul respectiv apare negru, iar dacă cele trei culori primare au valoarea 255, pixelul are culoarea albă.

• LUMINOZITATE ŞI CONTRAST RGB

Spre deosebire de cazul imaginilor monocrome, pentru creşterea calității imaginilor color poate fi necesară creşterea **echilibrului de culoare** sau a **contrastului de culoare**. Aceasta este o sarcină mai dificilă decât în cazul monocrom, nu numai datorită dimensiunilor suplimentare ale datelor, dar şi complexității percepției de culoare.

Modificările de luminozitate şi/sau contrast pot fi operate separat pe fiecare coordonată de culoare în parte, după principiile referitoare la imagini cu nuanțe de gri, discutate în capitolul 2, subsecțiunile 2.3.1 şi 2.3.2. Altfel spus, imaginea corespunzătoare fiecărei coordonate (fiecărei culori primare) este îmbunătățită printr-un algoritm de tip monocrom (conceput ca pentru o imagine cu niveluri de gri), după care imaginea se recompune după cele trei culori primare. Coordonatele de culoare ale fiecărui pixel din imaginea de intrare sunt transformate separat într-un alt set de coordonate de culoare. Este evident că, operând pe trei coordonate, (cum este cazul imaginilor color) îmbunătățirile aduse imaginilor pot fi mult mai eficiente din punctul de vedere al procesării ulterioare, decât în cazul unei singure coordonate (cum este cazul imaginilor monocrome).

• SEGMENTARE RGB

Principiul de realizare a segmentării în cazul imaginilor monocrome se generalizează în cazul imaginilor color. În această situație, segmentarea se bazează pe valorile celor trei culori primare, fiecare pixel fiind caracterizat printr-o **funcție nivel de culoare** (respectiv nivel de roşu, nivel de verde, nivel de albastru).

Segmentarea RGB se realizează relativ uşor în condițiile în care există o diferență importantă pe cel puțin una dintre componentele de culoare ale regiunii de interes şi, respectiv, regiunilor adiacente. Pentru acuratețea segmentării trebuie determinate, cu foarte mare precizie, intervalele din domeniul [0, 255] (pentru roşu, verde şi albastru) în care se situează valorile pixelilor. Modul de operare utilizează o valoare de prag minimă şi, respectiv, maximă, pentru fiecare din cele trei culori, ceea ce revine la exploatarea formulei (2.4.2) în sensul:

$$y=\begin{cases}1\,(\text{alb}), \text{ pentru } L_R \le x_R \le H_R \text{ si } L_G \le x_G \le H_G \text{ si } L_B \le x_B \le H_B \\ 0\,(\text{negru}), \text{ in rest}\end{cases}$$

Acest tip de segmentare este extrem de util în morfometria histologică şi histopatologică, deoarece structurilor de un anumit tip li se poate atribui un domeniu descris cu mai multă acuratețe prin cele trei condiții formulate mai sus (spre deosebire de o singură condiție, adică (2.4.2), în cazul imaginilor monocrome). Astfel, prin alegerea valorilor prag, şi, implicit, prin delimitarea unui interval specific pentru fiecare culoare (între 0 şi 255), se stabileşte care dintre elementele morfologice din imagine, corespunzătoare unui anumit domeniu RGB, vor fi păstrate sau nu în imaginea rezultată. Utilitatea segmentării RGB este demonstrată de studiile de caz din capitolul 4, secțiunile 4.5 şi 4.6.

Scanned with OKEN Scanner

Scanned with OKEN Scanner

ANEXA II

TEXTE SURSĂ ŞI SCHEME LOGICE

În cele cinci anexe care urmează sunt prezentate textele sursă ale macrourilor dezvoltate în limbaj KLIC, pentru studiile de caz din capitolul 4.

Textele conţin toate comentariile necesare (linii marcate cu simbolul #) pentru a asigura înţelegerea modului de programare. În aceste comentarii, fiecare imagine din galeria de imagini construită prin rularea unui anumit macrou este referită drept imgNumăr, unde Număr este numărul imaginii din galeria respectivă. De asemenea, comentariile macrourilor fac trimiteri sub forma (fig. 4.x.y) la toate figurile şi tabelele utilizate în descrierea studiilor de caz. Este evident că, pentru detaliile de sintaxă privind fiecare instrucţiune în parte, cititorul trebuie să apeleze la manualul de utilizare al Zeiss KS400.

Limbajul KLIC operează cu instrucţiuni (comenzi) scrise pe o singură linie, care se prelungeşte la numărul necesar de caractere, fără a se permite continuarea pe următoarea linie. O linie de comandă lungă poate fi citită pe monitor pe porţiuni, prin deplasarea cursorului la dreapta. Pentru reproducerea textului sursă din anexe, liniile de comandă lungi au fost fragmentate, punând între fragmente mesajul "**...continuare**". Cititorul interesat în a utiliza textul sursă va trebui să refacă sub KLIC continuitatea liniei de comandă.

Exceptând comenzile specifice operării cu imagini, sintaxa instrucţiunilor de uz comun din KLIC (scriere, citire pe terminal, operaţii aritmetice, structuri de control logic etc.) prezintă multiple similitudini cu convenţiile de scriere pseudocod şi, drept urmare, este uşor de urmărit de către un cititor cu experienţă minimă în programare. Cu toate acestea, în scopul creşterii comprehensibilităţii textelor sursă, am asociat şi schemele logice (*eng.* flowchart) corespunzătoare. În elaborarea lor s-a urmărit o descriere a procedurilor care să fie cât mai puţin dependentă de filozofia programării sub KLIC. *Considerăm că, atunci când se intenţionează transpunerea rutinelor KLIC în alte medii de programare ce permit analiza de imagine, schemele logice oferă un plus de informaţii ce complementează textele sursă şi facilitează reconstrucţia raţionamentelor ce stau la baza implementărilor prezentate în aceste anexe.*

Organizarea anexelor a fost gândită astfel încât fiecare schemă logică să fie plasată în imediata continuare a textului sursă aferent.

ANEXA II.1

- **MACROUL ODO_P1**

```
# ********************************************************************************************
# NUME: ODO_P1
# OBIECTIV: marcarea interactiva a contururilor celulare si nucleare a odontoblastelor, in vederea procesarii
# sub macroul ODO_P2
# APELEAZA: fisierul imagine "odonto.tif"
# RETURNEAZA: fisierele de imagine od1_des.tif, od2_des.tif, od3_des.tif, od4_des.tif
# METODA: utilizarea editorului grafic pentru marcarea interactiva a suprafetelor selectate (celule, nuclei) si
# operatii cu planul grafic
# ********************************************************************************************

#-------------------------------------------------------------------------------------------------------
# INITIALIZARE IMAGINI, PLAN GRAFIC SI FEREASTRA MESSAGES
#-------------------------------------------------------------------------------------------------------
# Initializare imagini
imgdelete "*"

 # Initializare plan grafic
Gclear 0

# Initializare fereastra Messages
showwindow "Messages", 1
write "@"
wait 10

# -------------------------------------------------------------------------------------------------------
# PROCESARE IMAGINI
# -------------------------------------------------------------------------------------------------------
# Incarcare imagine initiala in img1 (fig. 4.2.1)
imgload "odonto.tif",1
imgdisplay 1
update

# Copiere imagine initiala intr-o imagine noua – img2 (fig. 4.2.1)
imgcopy 1,2

# Pregatire img3 si img4 ca imagini binare (fundal negru) cu urmatoarea secventa de instructiuni
imgnew 3,512,512,1,"Grey"
imgnew 4,512,512,1,"Grey"
update

# Mod de lucru interactiv; se lanseaza Editorul grafic (fig. 4.2.2) pentru desenarea contururilor
# Desenare contururi pentru celule
imgdisplay 1
showwindow "Editor",1

# Secventa de operatii: 1) apare mesajul; 2) se marcheaza creionul si culoarea; 3) se deseneaza contururile;
# 4) se deselecteaza editorul (creionul in afara figurii); 5) se minimizeaza img1; 6) se reactiveaza fereastra
# de mesaj in care se apasa OK. Editorul ramane deschis; desenul se realizeaza in planul grafic
MBok "Se traseaza contururi in jurul celulelor"

# Contururile desenate in planul grafic se aplica pe img1 (cu culoarea cu care au fost desenate)
Gmerge 1,255
```

Scanned with OKEN Scanner

```
# Contururile desenate in planul grafic se aplica pe imaginea din img3 (cu culoare alba) (fig. 4.2.3)
Gmerge 3,255

# Planul grafic devine transparent
Gclear 0

# Desenare contururi pentru nuclei
imgdisplay 1

# Secventa de operatii: 1) apare mesajul; 2) se marcheaza creionul si culoarea; 3) se deseneaza contururile;
# 4) se deselecteaza editorul (creionul in afara figurii); 5) se minimizeaza img1; 6) se reactiveaza fereastra
# de mesaj in care se apasa OK. Editorul este deja deschis si se va inchide odata cu terminarea desenului;
# desenul se realizeaza in planul grafic
MBok "Se traseaza contururi in jurul nucleilor"

# Inchidere Editor grafic
showwindow "Editor",0

# Contururile desenate in planul grafic se aplica pe img1 (cu culoarea cu care au fost desenate) (fig. 4.2.5)
Gmerge 1,255

# Contururile din planul grafic se aplica pe img4 (cu culoare alba) (fig. 4.2.4)
Gmerge 4,255

# Planul grafic devine transparent
Gclear 0

# Imaginile img1, img2, img3, img4 (existente in galerie), anterior create, sunt salvate in fisiere de imagine (.tif)
# destinate a fi procesate in macroul ODO_P2 in care se efectueaza numai masuratori (fara lucru interactiv)
imgsave 1, "od1_des.tif"
imgsave 2, "od2_des.tif"
imgsave 3, "od3_des.tif"
imgsave 4, "od4_des.tif"

beep 200,440
pause "FINAL ODO_P1",0
stop
```

- **MACROUL ODO_P2**

```
# ********************************************************************************************
# NUME: ODO_P2
# OBIECTIV: determinarea dimensiunilor odontoblastelor (arie celulara, arie nucleara, raport nucleo-citoplasmatic)
# APELEAZA: fisierele imagine od1_des.tif, od2_des.tif, od3_des.tif, od4_des.tif
# RETURNEAZA: informatii numerice organizate tabelar si reprezentari grafice de tip histograma si diagrama
# de dispersie, pentru ariile masurate
# BAZE DE DATE: odonto.d00 (creare si utilizare)
# METODA: masurarea automata a trasaturilor ce caracterizeaza suprafetele selectate interactiv
# (ca regiuni ale campului)
# ********************************************************************************************

# ---------------------------------------------------------------------------------------------
# INITIALIZARE IMAGINI, PLAN GRAFIC SI FEREASTRA MESSAGES
# ---------------------------------------------------------------------------------------------
# Initializare imagini
imgdelete "*"
```

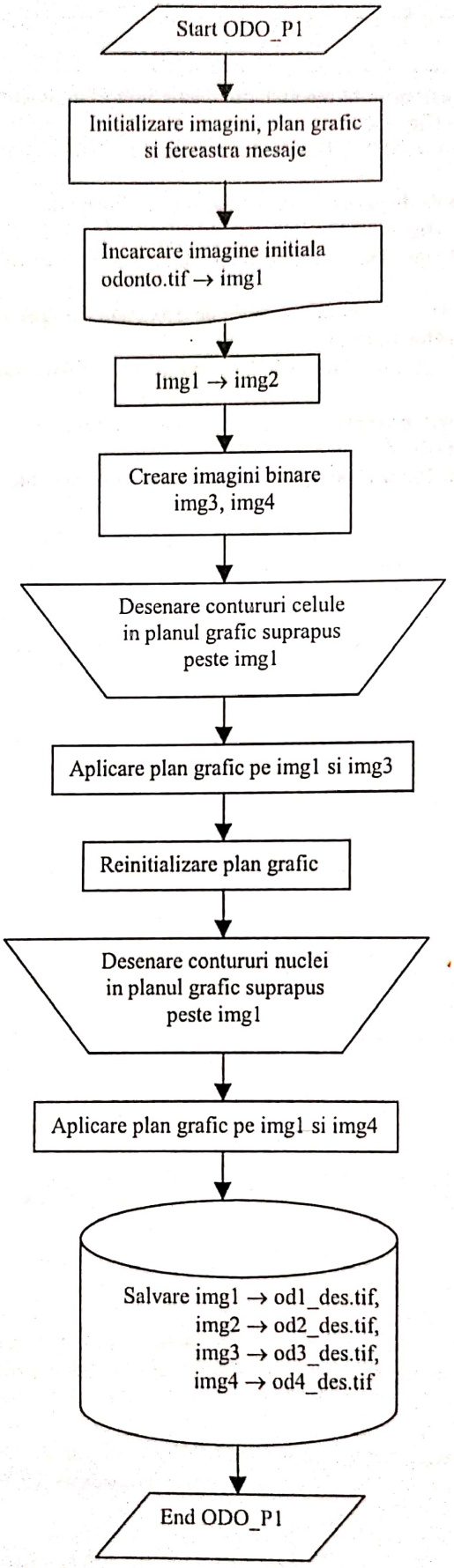
Scanned with OKEN Scanner

```
# Initializare plan grafic
Gclear 0

# Incarcare imagini rezultate din ODO_P1.mcr si creare de imagini binare necesare masuratorilor
# Incarcare imagine color cu contururile celulare si nucleare (fig. 4.2.5)
imgload "od1_des.tif", 1
imgdisplay 1

# Incarcare imagine originala (fara contururi) (fig. 4.2.1)
imgload "od2_des.tif", 2

# Incarcare imagine binara cu contururile celulare (fig. 4.2.3)
imgload "od3_des.tif", 3

# Incarcare imagine binara cu contururile nucleare (fig. 4.2.4)
imgload "od4_des.tif", 4

# ---------- ----------
# PROCESARE IMAGINI
# ----------
# Umplere contururilor celulare din img3 si depozitare in img5 (fig. 4.2.6)
binfill 3,5

# Umplere contururilor nucleare din img4 si depozitare in img6 (fig. 4.2.7)
binfill 4,6

# Suprafetele citoplasmatice (albe) din img7 (fig. 4.2.8) rezulta ca SAU-EXCLUSIV intre suprafetele celulare
# din img5 si suprafetele nucleare din img6
binxor 5,6,7

# ----------
# ANALIZA CANTITATIVA
# ----------
# Declarare si definire trasaturi care vor face obiectul masurarii
# FLDCOUNT (predefinita) de tip FIELD – masoara numarul [#] de regiuni din camp
# CellArea (definita de utilizator) de tip regiune – masoara aria fiecarei celule marcate
# NcArea (definita de utilizator) de tip regiune – masoara aria fiecarui nucleu marcat
# NcRatio (definita de utilizator) de tip regiune – calculeaza raportul arie nucleu / arie citoplasma
# Incarcare fisier de proprietati default.ms
MSload "default"

# Definire trasaturi
MSsetprop "REGIONFEAT","CellArea[UNIT^2]=AREAF,NcArea[UNIT^2]=AREAF-AREA, ...continuare
NcRatio[%]=(AREAF-AREA)*100/AREAF"

# Declarare nume baza de date
DBname="odonto"

# Masurare trasaturi de tip regiune declarate si definite anterior pe imaginea originala din img2 cu
# imaginea-masca din img7; pentru memorarea rezultatelor se utilizeaza aceeasi baza de date ca mai sus
MSmeasmask 7,2,DBname,0,1,10

# Afisare tabel cu informatii din baza de date (fig. 4.2.9) cu rezultatele masuratorilor organizate matriceal:
# linii – regiuni; coloane – CellArea, NcArea, NcRatio
datalist DBname

pause "Tabel – masuratorile pe regiuni",0
```

Scanned with OKEN Scanner

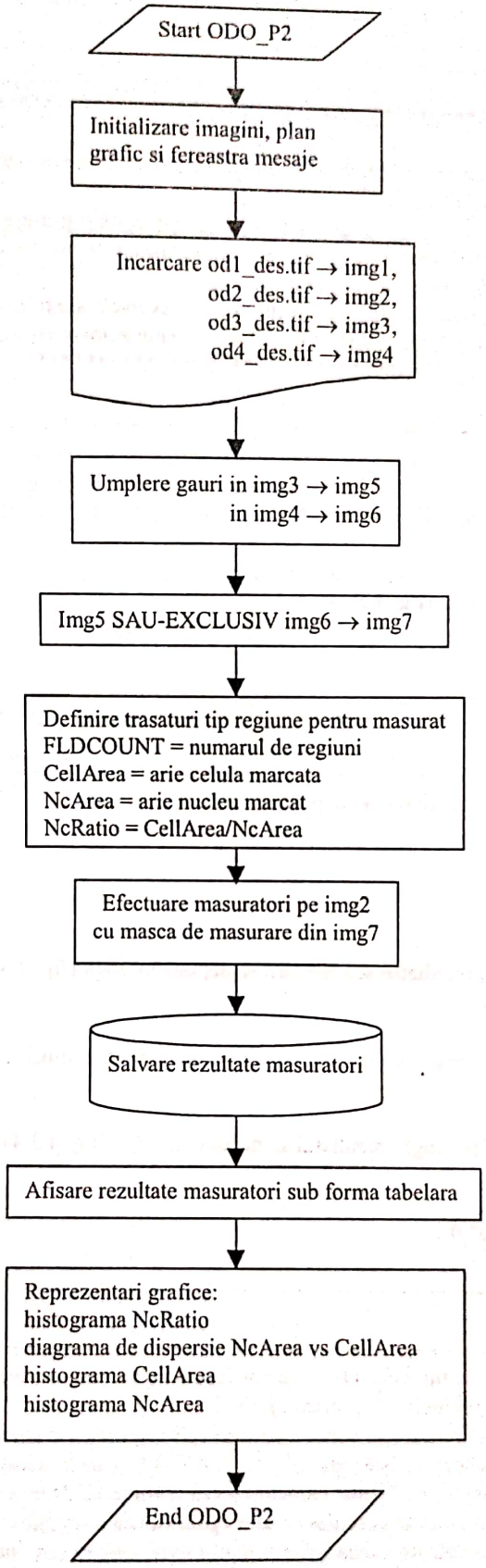
```
# ------------------------------------------------------------------------------------------------------------------------
# REPREZENTARI GRAFICE
# ------------------------------------------------------------------------------------------------------------------------
# Reprezentare grafica (histograma): procent din numarul de masuratori in functie de valorile NcRatio;
# fereastra grafica permite scalari (fig. 4.2.10)
datahisto DBname,"NcRatio",0,14,0.,70.0,8,"Distribution of NcArea/CellArea ratios"

# Reprezentare grafica (diagrama de dispersie): valorile NcArea in functie de valorile CellArea;
# fereastra grafica permite scalari (fig. 4.2.11)
datascat DBname,"CellArea",DBname,"NcArea",0,0.,4500.,0.,1500.,"NcArea versus CellArea"

# Reprezentare grafica (histograma): procent din numarul de masuratori in functie de valorile CellArea;
# fereastra grafica permite scalari (fig. 4.2.12)
datahisto DBname,"CellArea",0,12,1250.0,4250.0,7,"Distribution of CellArea values"

# Reprezentare grafica (histograma): procent din numarul de masuratori in functie de valorile NcArea;
# fereastra grafica permite scalari (fig. 4.2.13)
datahisto DBname,"NcArea",0,10,250.0,1500.0,7,"Distribution of NcArea values"

beep 200,440
pause "FINAL ODO_P2",0
stop
```

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **ODO_P1**.

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **ODO_P2**.

Scanned with OKEN Scanner

ANEXA II.2

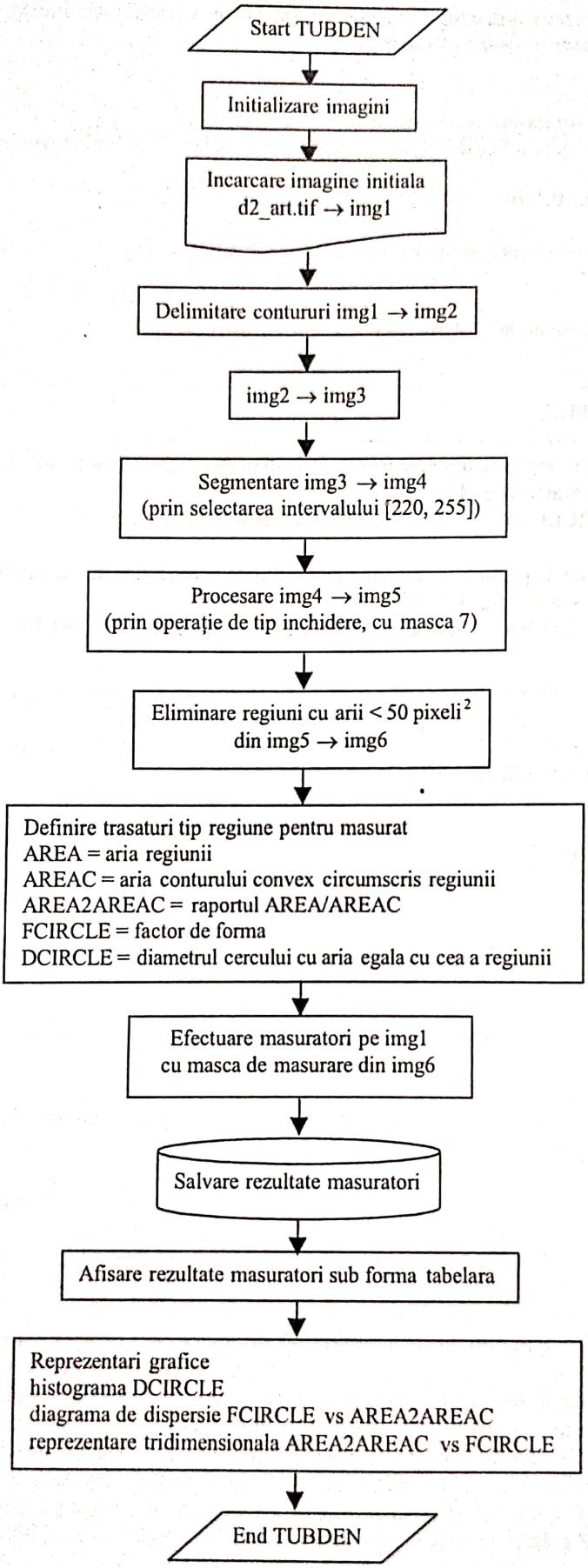
- **MACROUL TUBDEN**

```
# ********************************************************************************************
# NUME: TUBDEN
# OBIECTIV: determinarea diametrelor tubulilor dentinari in sectiune transversala
# APELEAZA: fisierul imagine "d2_art.tif"
# RETURNEAZA: informatii numerice organizate tabelar si reprezentari grafice de tip histograma, diagrama de
# dispersie si grafic tridimensional pentru numarul diametrelor masurate functie de valorile acestora
# BAZE DE DATE: tub_den.d00 (creare si utilizare)
# METODA: se delimiteaza regiunile definite de sectiunile transversale ale tubulilor, se analizeaza circularitatea
# acestora si se determina valorile diametrelor echivalente corespunzatoare regiunilor respective
# ********************************************************************************************

#-----------------------------------------------------------------------------------------------------------
# INITIALIZARE IMAGINI
# -----------------------------------------------------------------------------------------------------------
# Initializare imagini
imgdelete "*"

# Incarcare imagine initiala in img1 (fig. 4.3.1)
imgload "d2_art.tif",1
imgdisplay 1

#-----------------------------------------------------------------------------------------------------------
# PROCESARE IMAGINI
#-----------------------------------------------------------------------------------------------------------
# Delimitare contururi pe img1; rezultatul se depune in img2 (fig. 4.3.2)
delin 1,2,1,17

# Copiere img2 in img3
imgcopy 2,3

# Segmentare aplicata pe img3; rezultatul segmentarii se depune in img4 (fig. 4.3.3)
dislev 3,4,220,255,1

# Dilatare urmata de erodare pe img4 cu masca 7; rezultatul se depune in img5
binclose 4,5,7,3

# Eliminare arii sub 50 pixel^2 pe img5; rezultatul se depune in img6 (fig. 4.3.4)
binscrap 5,6,0,50,0

pause "imagine binara pregatita",0

# -----------------------------------------------------------------------------------------------------------
# ANALIZA CANTITATIVA
# -----------------------------------------------------------------------------------------------------------
# Declarare si definire trasaturi de tip REGION care vor face obiectul masuratorilor
# AREA (predefinita) este aria regiunii – masurata in pixel^2
# AREAC (predefinita) este aria conturului convex circumscris regiunii – masurata in pixel^2
# AREA2AREAC (definita de utilizator) este raportul AREA/AREAC; caracterizeaza apropierea de un contur convex
# FCIRCLE (predefinita) este factor de forma; caracterizeaza apropierea de un contur circular
# DCIRCLE (predefinita) este diametrul cercului cu aria egala cu cea a regiunii – masurat in pixel
# Daca in directorul Ks400\Conf\Ms nu exista fisierul numit user, care sa contina exact declararea si definirea
# acestor trasaturi, atunci aceste trasaturi se pot introduce interactiv, prin cutia de dialog Set Features
MSsetfeat "REGIONFEAT"
```

Scanned with OKEN Scanner

```
# Trasaturile introduse interactiv prin cutia de dialog sunt salvate intr-un fisier din directorul
# Ks400\Conf\Ms, cu ajutorul comenzii MSsave
MSsave "user"

# Efectuare masuratori pentru trasaturile declarate
# Rezultatele masuratorilor pentru fiecare regiune in parte constituie cate o linie in organizarea matriceala
# a bazei de date "tub_den"
MSmeasmask 6,1,"tub_den",0,1,10

# Afisare tabel cu informatiile numerice din baza de date tub_den (fig. 4.3.5)
datalist "tub_den",0,0

pause "Tabelul cu informatiile numerice",0

# ----------------------------------------------------------------------------------------------------------------------
# REPREZENTARI GRAFICE
# ----------------------------------------------------------------------------------------------------------------------
# Histograma de frecvenţe relative: % din numarul total de diametre (regiuni) in functie de valoarea diametrelor;
# fereastra grafica permite scalari (fig. 4.3.6)
datahisto "tub_den","DCIRCLE",0,10,16,36.,20.,"Distribution of equivalent diameters"

# Scattergrama (diagrama de dispersie): dependenta factorului de convexitate de factorul de forma;
# fereastra grafica permite scalari (fig. 4.3.7)
datascat "tub_den","AREA2AREAC","tub_den","FCIRCLE",0,0.6,1.,0.6,1.,"FCIRCLE versus AREA2AREAC"

# Distributie (reprezentare tridimensionala): numar de diametre masurate ca functie de cei doi factori (variabilele
# AREA2AREAC si FCIRCLE); fereastra grafica permite scalari (fig. 4.3.8)
datadist "tub_den","AREA2AREAC","tub_den","FCIRCLE",0,8,8,0.6,1.,0.6,1., ...continuare
"Image of correlation between AREA2AREAC and FCIRCLE"

beep 200,440
pause "FINAL TUBDEN",0
stop
```

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **TUBDEN**.

Scanned with OKEN Scanner

ANEXA II.3

- **MACROUL DENTIN**

```
# ********************************************************************************************
# NUME: DENTIN
# OBIECTIV: determinarea diametrelor tubulilor dentinari in sectiune longitudinala
# APELEAZA: fisierul imagine "denti_n.tif"
# RETURNEAZA: informatii numerice organizate tabelar si reprezentari grafice de tip histograma, pentru
# numarul diametrelor masurate functie de valorile acestora;
# in final, se calculeaza si valoarea medie a tuturor diametrelor detectabile in imagine
# BAZE DE DATE: dentin.d00, dentin1.d00 (creare si utilizare)
# METODA: detectarea diametrelor se realizeaza prin parcurgerea orizontala, de la stanga la dreapta, a imaginii
# binare rezultate in urma segmentarii si indepartarii regiunilor nesemnificative; algoritmul se bazeaza pe
# identificarea frontierelor alb-negru, negru-alb si reducerea iterativa a zonelor albe prin erodare laterala
# ********************************************************************************************

# ----------------------------------------------------------------------------------------------------
# INITIALIZARE IMAGINI, PLAN GRAFIC SI FEREASTRA MESSAGES
# ----------------------------------------------------------------------------------------------------
# Initializare imagini, pregatire plan grafic
imgdelete "*"
Gclear 0
update

# Pregatire fereastra Messages
showwindow "Messages", 1
write "@"
wait 100

# Pregatire bordura pentru imaginile binare pe care se vor efectua masuratorile
# Creare imagine noua in img7
imgnew 7,512,512,1,"Grey"
imgdisplay 7

# Creare dreptunghi cu bordura de 12 pixeli in planul grafic
Grectg 12,12,488,488,11,7

# Copiere dreptunghi din planul grafic in img7 si setare dreptunghi pe culoare alba
Gmerge 7,255

# Planul grafic devine transparent
Gclear 0

# Incarcare imagine initiala in img1 (fig. 4.4.1)
imgload "denti_n.tif",1
update

# ----------------------------------------------------------------------------------------------------
# PROCESARE IMAGINI
# ----------------------------------------------------------------------------------------------------
# Segmentare aplicata pe img1 si depunere in img2 (fig. 4.4.2)
dislev 1,2,165,255,1
update

# Deschidere verticala cu masca 3, aplicata consecutiv de 10 ori, urmata de cate o deschidere cu masca 1 si 7;
# rezultatul se depune in img3 (fig. 4.4.3)
binopen 2,3,3,10
```

Scanned with OKEN Scanner

```
binopen 3,3,1,1
binopen 3,3,7,1

# Eliminare arii sub 500 pixel^2 din img3; rezultatul se depune in img4 (fig. 4.4.4)
binscrap 3,4,1,500,0

# Umplere gauri pe img4 cu rezultat in img5
binfill 4,5

# Închidere cu masca 7 aplicata pe img5; rezultatul se depune in img5 (fig. 4.4.5)
binclose 5,5,7,1

# Operatie booleana ŞI cu img7 pentru introducerea bordurii, cu rezultat in img5 (fig. 4.4.6)
binand 5,7,5

# Img5 se copie in img6; in erodarea succesiva se va lucra cu ambele imagini
imgcopy 5,6

# In img8 se pastreaza imaginea binara initiala din img5, neafectata de erodari, pe care se va aplica
# FLDMEANCHORD
imgcopy 5,8

pause "imagini binare pregatite",0

# ----------------------------------------------------------------------------------------------------------------------
# ANALIZA CANTITATIVA
# ----------------------------------------------------------------------------------------------------------------------
# Declarare trasaturi de tip FIELD cu care urmeaza a se opera
# Declararea se poate realiza si interactiv folosind MSsetfeat "FIELDFEAT" care deschide o cutie de dialog
# Unitatea de masura conventionala pentru lungime are semnificatia de pixel
MSsetprop "UNIT","pixel"

# FLDINTERCEPT este un numar [#] (este predefinita)
MSsetprop "FIELDFEAT","FLDINTERCEPT[#]"

# DIFF_FLDINTER este un numar [#] egal cu diferenta dintre FLDINTERCEPT la doua iteratii succesive;
# se va defini de utilizator; pana la definire are valoarea 1
MSsetprop "FIELDFEAT",",,DIFF_FLDINTER[#]=1"

# DIAM_COUNT este un numar [#] egal cu DIFF_FLDINTER / 2; se va defini de utilizator;
# pana la definire are valoarea 1
MSsetprop "FIELDFEAT",",,DIAM_COUNT[#]=1"

# DIAMETER este o lungime [UNIT] egal cu index linie * 2 – 2; se va defini de utilizator;
# pana la definire are valoarea 1
MSsetprop "FIELDFEAT",",,DIAMETER[UNIT]=1"

# Structura while pentru completarea campului (coloanei) FLDINTERCEPT
# Masurare interceptari prin erodare
# Erodare cu masca 1 (orizontala) pana cand numarul de interceptari alb-negru devine 0, adica pana la epuizarea,
# prin erodare, a tuturor zonelor albe
# Efectul erodarii consta in eliminarea pixelilor extremi de pe fiecare segment orizontal din zonele albe
# Step memoreaza a cata masuratoare de interceptari se realizeaza
# Pentru ultima valoare a lui step, FLDINTERCEPT=0
step:=1

# Controlare scriere in baza de date (0 = la inceput)
app:=0
```

Scanned with OKEN Scanner

```
# Contorizare numar de interceptari
ix:=1

while (ix >=1)
# Efectuare masuratori pentru trasatura FLDINTERCEPT (predefinita)
        MSmeasmask 6,1,"dentin",app,2,10
        DBfirstline "dentin"
        step1=step-1

# Pregatire pozitionare pentru scriere in baza de date (intre doua apeluri ale lui MSmeasmask, nu se pastreaza
# pozitionarea scrierii)
# La prima scriere step1=0 si secventa se sare
        if (step1>0)
                for i=1, i<=step1, i=i+1
                        DBnextline "dentin"
                endfor
        endif

# Extragere numar de interceptari (care asigura si controlul buclei while)
        DBgetvalue "dentin","FLDINTERCEPT",x
        ix=int (x)

# Scriere de control in fereastra Messages
        write ix

# Erodare cu cate un pixel la extremitatea fiecarui segment orizontal din zona alba
        binerode 6,5,1,1
        imgcopy 5,6

# Permite vizualizarea in fereastra a erodarii pas cu pas prin actualizarea imaginii in fiecare iteratie (fig. 4.4.7)
        update

# Se asigura ca in iteratia viitoare scrierea in baza de date se va realiza la sfarsit
        app:=1
        step=step+1
endwhile

# Afisare tabel cu informatii numerice numai cu FLDINTERCEPT completat
datalist "dentin",0,1
pause "coloana 1",0

# Structura for pentru completarea campului (coloanei) DIFF_FLDINTERCEPT
# Se utilizeaza variabila dia
# Calculare diferenta dintre numarul de interceptari corespunzand la doua erodari succesive dia=x1-x2
# x1 > x2 memoreaza FLDINTERCEPT pentru linia anterioara si linia curenta
DBfirstline "dentin"

# Extragere din baza de date a valorii trasaturii FLDINTERCEPT; se atribuie lui x1
DBgetvalue "dentin","FLDINTERCEPT",x1

# Initializare variabila in care se calculeaza diferenta
dia:=0.0

# Prima data se seteaza valoarea 0 (neavand masuratoare anterioara)
DBsetvalue "dentin","DIFF_FLDINTER",dia

# Ultima evaluare se realizeaza pe linia ultima (step)
# Pornind de la prima linie, ultima pozitionare se realizeaza la step1=step-1
for i=1; i <= step1; i = i+1
```

Scanned with OKEN Scanner

```
    # Masurare trasatura DIFF_FLDINTERCEPT (definita de utilizator)
            DBnextline "dentin"
            DBgetvalue "dentin","FLDINTERCEPT",x2
            dia=x1-x2

    # Scriere in campul (coloana) DIFF_FLDINTER din baza de date
            DBsetvalue "dentin","DIFF_FLDINTER",dia
            x1=x2
    endfor

    # Afisare datalist cu FLDINTERCEPT, DIFF_FLDINTER completate
    datalist "dentin",0,1
    pause "coloanele 1,2",0

    # Structura for pentru completarea campurilor (coloanelor) DIAM_COUNT, DIAMETER
    # Utilizare variabile fra, stepby2
    # Numarul de diametre (masurate pe zona alba) este dat de jumatatea numarului de interceptari fra=dia/2
    # Valoarea diametrului este data de dublul indicelui liniei din care se scad cei doi pixeli extremi erodati
    # anterior stepby2=stepf*2-2
    total1=0.
    total2=0.
    stepfl:=1.0
    DBfirstline "dentin"

    # Asigurare de atatea iteratii cate linii au rezultat de la FLDINTERCEPT (step)
    for i=1; i <= step1+1; i = i+1
            DBgetvalue "dentin","DIFF_FLDINTER",dia
            fra:=dia/2
            stepby2:=stepfl*2-2
            DBsetvalue "dentin","DIAM_COUNT",fra
            DBsetvalue "dentin","DIAMETER",stepby2
            DBnextline "dentin"

    # Suma tuturor diametrelor masurate
            total1=total1+fra*stepby2

    # Numarul tuturor diametrelor masurate
            total2=total2+fra
            stepfl=stepfl+1
    endfor

    # Calculare diametru mediu
    mea_dia=total1/total2

    # Scriere in fereastra Messages
    write total1
    write total2
    write mea_dia

    # Afisare tabel cu informatii numerice continand FLDINTERCEPT, DIFF_FLDINTER, DIAM_COUNT,
    # DIAMETER completate (fig. 4.4.8)
    datalist "dentin",0,1
    pause "coloanele 1,2,3,4",0

    # ----------------------------------------------------------------------------------------------------
    # REPREZENTARI GRAFICE
    # ----------------------------------------------------------------------------------------------------
    # Reprezentare grafica pentru tot domeniul de valori ale diametrelor; fereastra grafica permite scalari (fig. 4.4.9)
    datacross "dentin","DIAMETER","DIAM_COUNT",0,36,1,0.,120.,0.,650.,...continuare
     "Distribution for the whole range of intercepted diameters"
```

Scanned with OKEN Scanner

```
pause "Grafic – tot domeniul de valori"

# ------------------------------------------------------------------------------------------
# VERIFICARE MASURATORI (ALTA METODA)
# ------------------------------------------------------------------------------------------
# Deschidere cutie de dialog Set Features, lucru interactiv
# FIELDFEAT este selectat automat si se seteaza un fisier (de exemplu "curent")
# Indepartare trasaturi existente si setare trasatura FLDMEANCHORD (predefinita), care se va masura
MSsetfeat "FIELDFEAT"
app:=0

# Masurare coarda medie din img8
MSmeasmask 8,1,"dentin1",app,2,10

# Afisare tabel cu informatii numerice pentru FLDMEANCHORD
datalist "dentin1",0,1
pause "Valoare FLDMEANCHORD"

beep 200,440
pause "FINAL DENTIN",0
stop
```

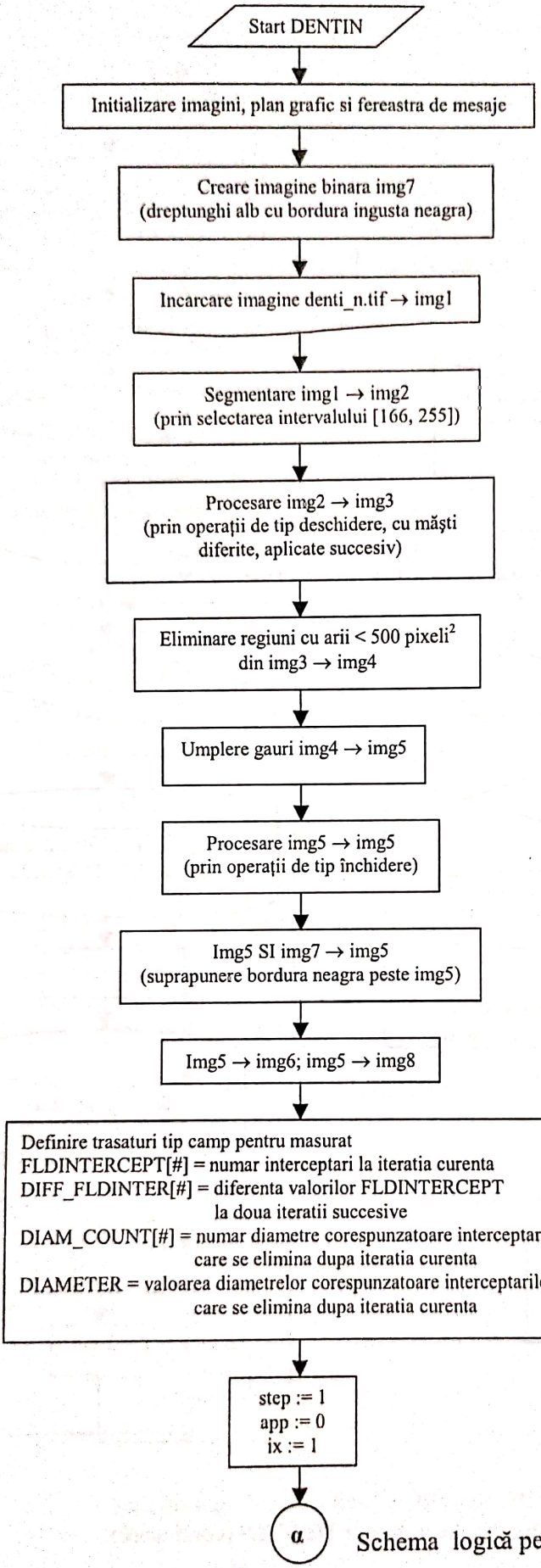
- **MACROUL DENT**

```
# ********************************************************************************************
# NUME: DENT
# OBIECTIV: studierea trasaturilor cantitative pentru submultimea de tubuli dentinari caracterizati prin diametre
# cuprinse intre 16 si 36 pixeli
# APELEAZA: baza de date dentin.d00
# RETURNEAZA: informatii numerice organizate tabelar si reprezentari grafice de tip histograma, pentru
# numarul diametrelor masurate functie de valorile acestora;
# domeniul de valori ale diametrelor este limitat la gama 16 – 36 pixeli constatata in sectiune transversala
# (via macroul TUBDEN)
# BAZE DE DATE: den16_36.d00 (creare si utilizare)
# ********************************************************************************************

# ------------------------------------------------------------------------------------------
# CREARE BAZA DE DATE NOUA DIN BAZA DE DATE dentin.d00
# ------------------------------------------------------------------------------------------
# Restrangere evaluari la domeniul de valori ale diametrelor din TUBDEN.
# Pozitionare pe linia a 9-a
DBfirstline "dentin"
for i=1;i<=8;i=i+1
        DBnextline "dentin"
endfor

# Creare baza de date noua, care include numai liniile aferente diametrelor din gama determinata de TUBDEN
DBnew "den16_36",2
DBsetcolumn "den16_36",1,"DIA_COUNT","float","#"
DBsetcolumn "den16_36",2,"DIAMET","float","pixel"

# Copiere din baza de date "dentin" in baza de date "den16_36"
DBfirstline "den16_36"
total1=0.
total2=0.
for i=1;i<=10;i=i+1
        DBgetvalue "dentin","DIAM_COUNT",fra
        DBaddline "den16_36"
```

Scanned with OKEN Scanner

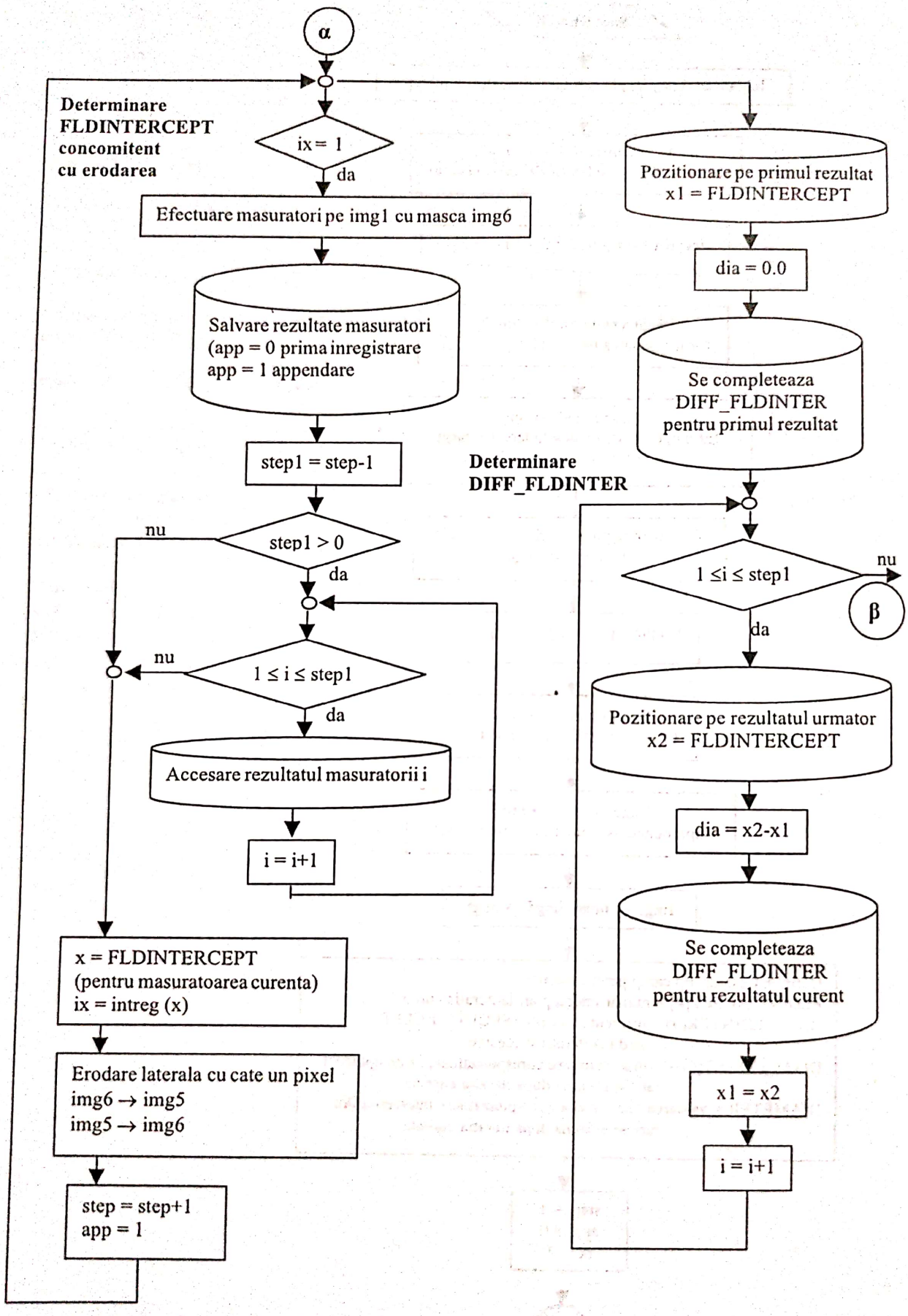
```
        DBsetvalue "den16_36","DIA_COUNT",fra
        DBgetvalue "dentin","DIAMETER",stepby2
        DBsetvalue "den16_36","DIAMET",stepby2
        DBnextline "dentin"
        total1=total1+fra*stepby2
        total2=total2+fra
endfor
mea_dia=total1/total2

# Scriere valori numerice in fereastra Messages
write total1
write total2
write mea_dia

# Afisare tabel informatii numerice pentru baza de date redusa la domeniul TUBDEN
datalist "den16_36",0,1

pause "Date – domeniul din TUBDEN",0

#-----------------------------------------------------------------------------------------------------------------------
# REPREZENTARI GRAFICE
#-----------------------------------------------------------------------------------------------------------------------
# Reprezentare grafica corespunzand domeniului TUBDEN; fereastra grafica permite scalari (fig. 4.4.10)
datacross "den16_36","DIAMET","DIA_COUNT",0,10,1,16.,36.,0.,700., ...continuare
"Distribution for intercepted diameters between 16-36 pixels"

beep 200,440
pause "FINAL DENT"
stop
```

Scanned with OKEN Scanner



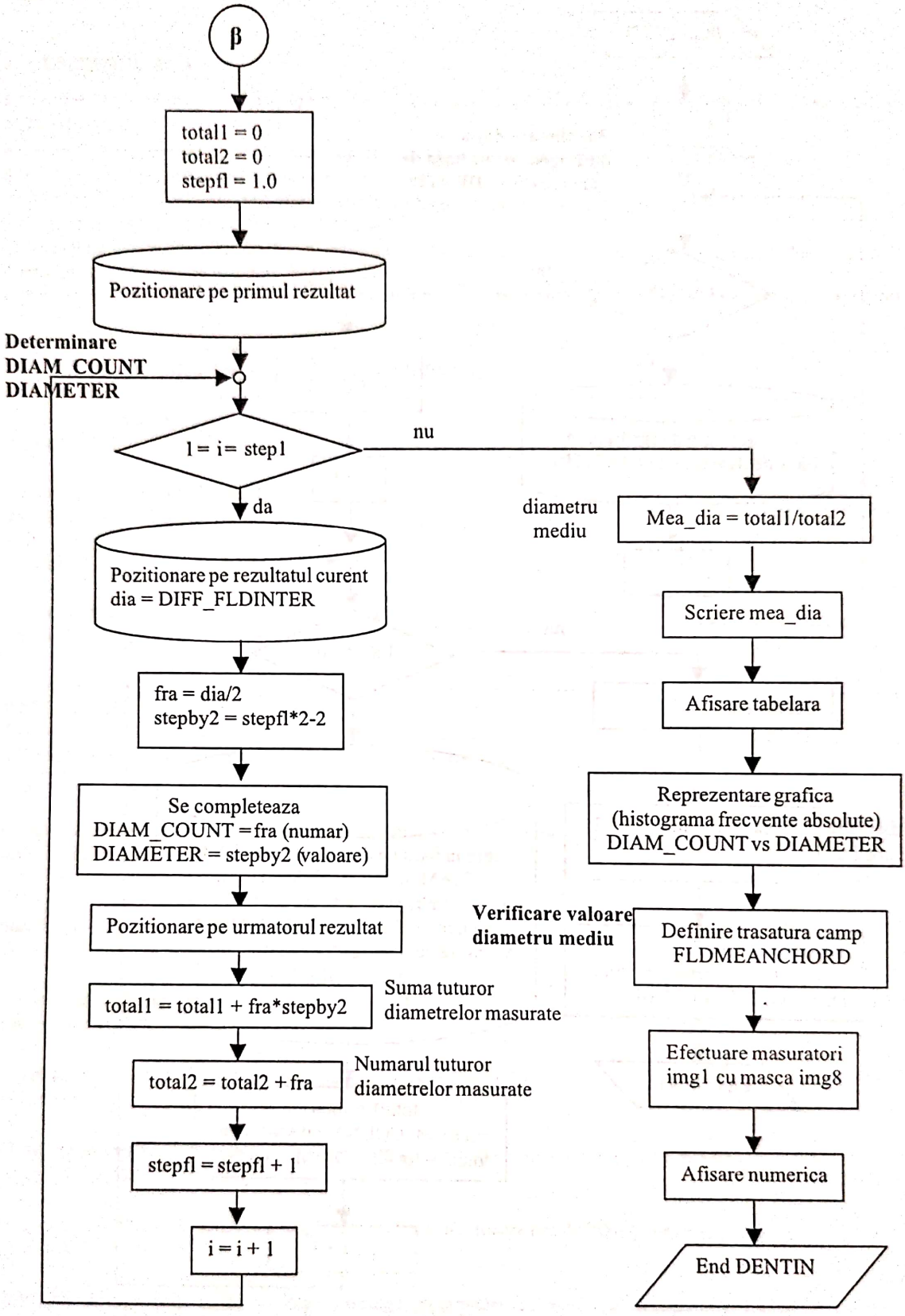
Schema logică pentru funcţia utilizator **DENTIN**.

Scanned with OKEN Scanner



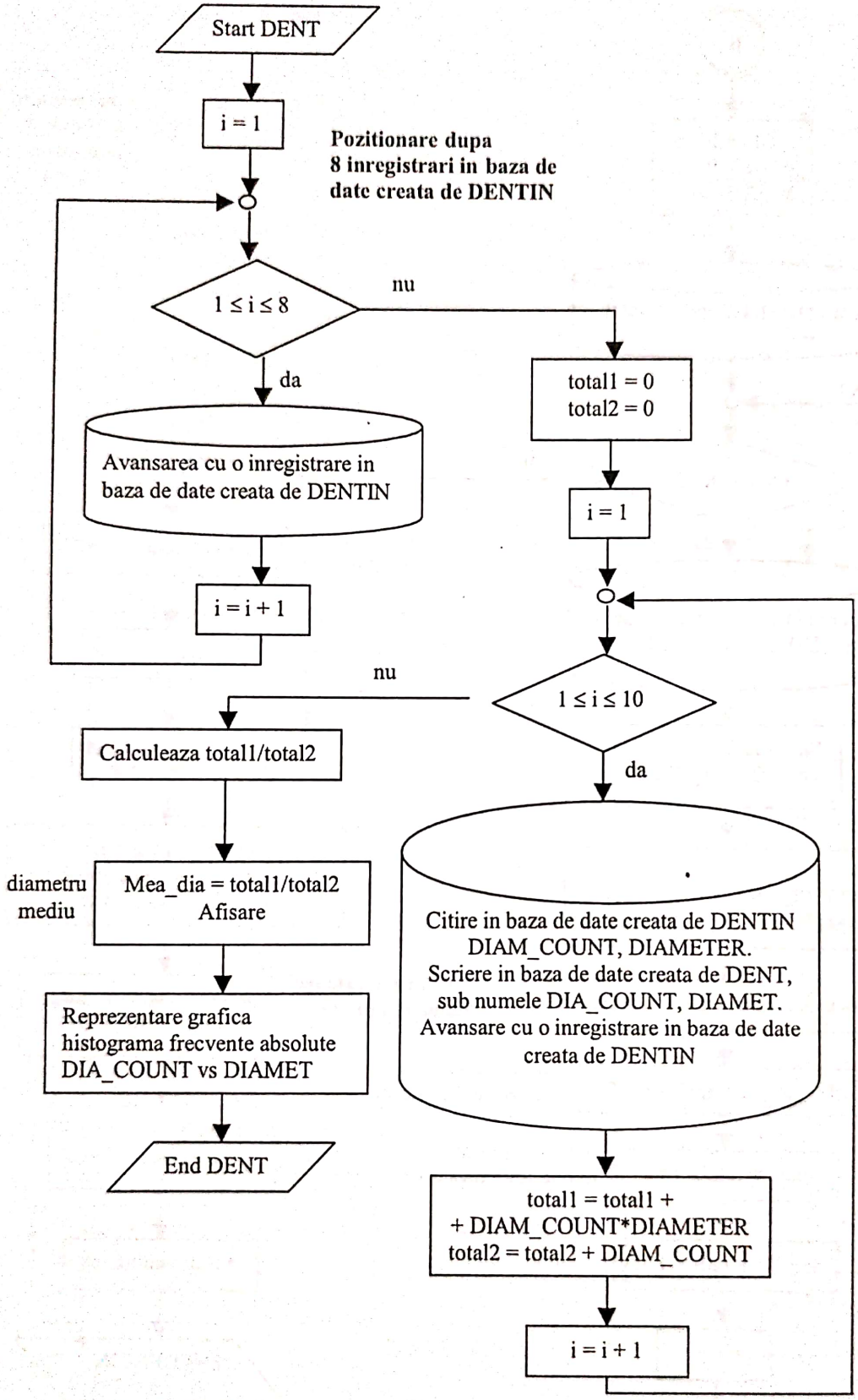
Schema logică pentru funcția utilizator **DENTIN** (continuare).

Scanned with OKEN Scanner



Schema logică pentru funcţia utilizator **DENTIN** (continuare).

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **DENT**.

Scanned with OKEN Scanner

ANEXA II.4

- **MACROUL ADK**

```
# ********************************************************************************************
# NUME: ADK
# OBIECTIV: determinarea ariilor ocupate de insulele tumorale si respectiv stroma adiacenta
# APELEAZA: fisierul imagine "adk.tif"
# RETURNEAZA: informatii numerice afisate in fereastra de mesaje
# BAZE DE DATE: adk.d00 (creare si utilizare)
# METODA: delimitarea regiunilor de interes prin segmentare si prin operatii specifice imaginilor binare;
# masurarea automata ca trasaturi de camp
# ********************************************************************************************

# ---------------------------------------------------------------------------------------------------------
# INITIALIZARE IMAGINI SI FEREASTRA MESSAGES
# ---------------------------------------------------------------------------------------------------------
# Initializare imagini
imgdelete "*"

# Initializare fereastra Messages
showwindow "Messages", 1
write "@"
wait 10

# Incarcare imagine initiala in img1 (fig. 4.5.1)
imgload "adk.tif",1

# ---------------------------------------------------------------------------------------------------------
# PROCESARE IMAGINI
# ---------------------------------------------------------------------------------------------------------
# Transformare imagine initiala color in imagine cu nuante de gri – img2
imgRGB2grey 1,2

# Segmentare interactiva a imaginii initiale – obtinere imagine binara – img3 (fig. 4.5.2)
dislev 2,3,255,255,1

# Umplere gauri din imaginea binara – img4 (fig. 4.5.3)
binfill 3,4


# Indepartare arii sub 7000 pixeli^2 – img5 (fig. 4.5.4)
binscrap 4,5,0,7000,0

# Operatie booleana NU pe img5 cu depunere in img6 (fig. 4.5.5)
binnot 5,6

# Segmentare interactiva a imaginii initiale – obtinere imagine binara – img7 (fig. 4.5.6)
dislev 2,7,0,90,1

# Operatie aritmetica de scadere a img7 din img5, cu depunerea rezultatului in img8 (fig. 4.5.7)
subtract 5,7,8,2

# Operatie aritmetica de scadere a img7 din img6, cu depunerea rezultatului in img9 (fig. 4.5.8)
subtract 6,7,9,2

pause
```

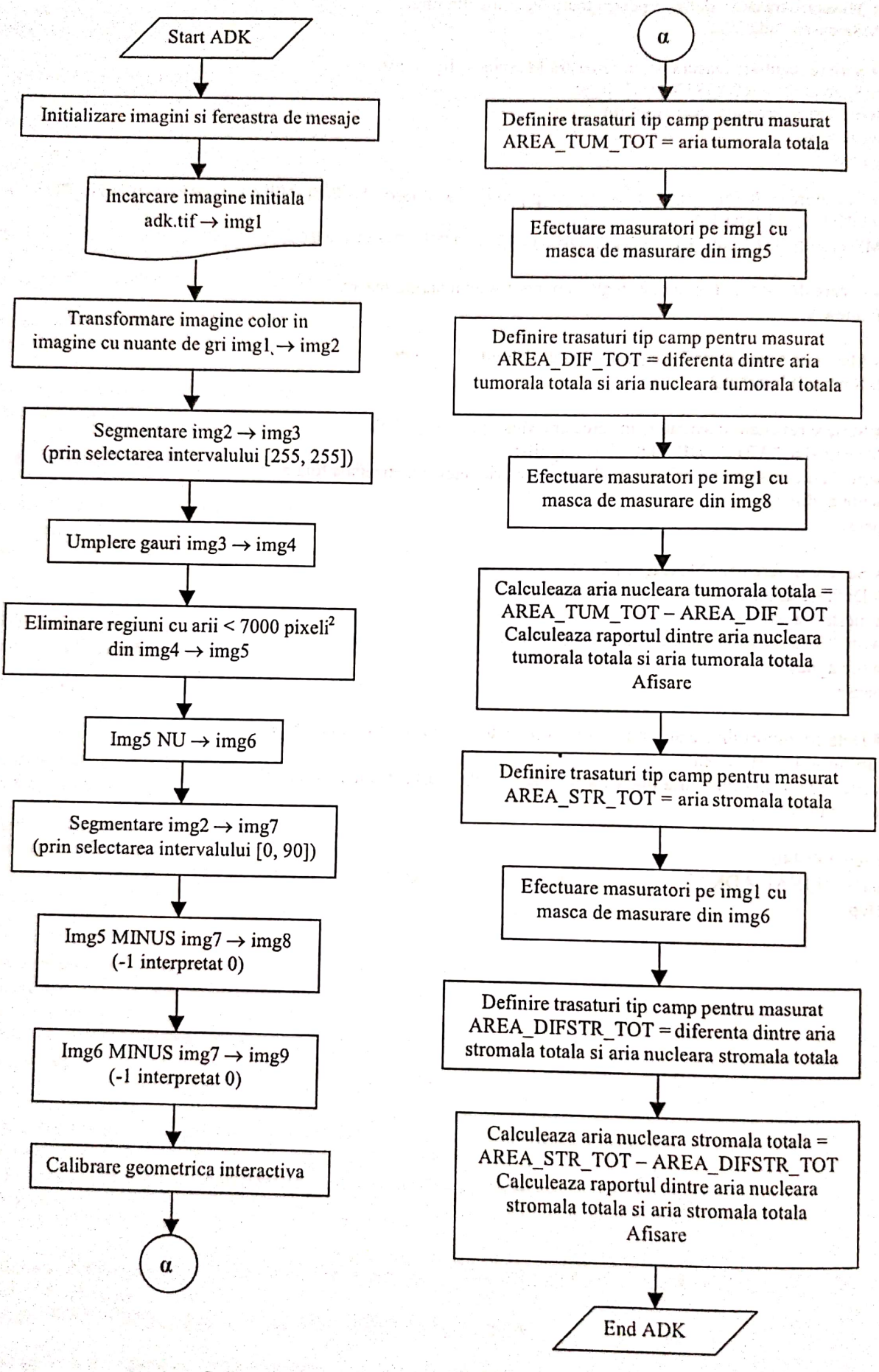
Scanned with OKEN Scanner

```
# ---------------------------------------------------------------------------------------------------------
# ANALIZA CANTITATIVA
# ---------------------------------------------------------------------------------------------------------
# Calibrare interactiva pentru masurare geometrica:
# distance X = 152000/400 = 380 microni, distance Y = 100000/400 = 250 microni
MSsetgeom

# Declarare nume pentru noua baza de date
DBnew "adk"

# Declarare si definire trasaturi de tip camp pentru masuratori: din FIELDFEAT – AREA_TUM_TOT...
# UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_TUM_TOT[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img5, care are rol de imagine-masca
RGnew 5,1

# Masurare trasaturi definite pentru toate regiunile din img5
MSmeasall "adk",0,2

# Scriere rezultate masuratori in fereastra Messages (fig. 4.5.9)
MSgetvalue "AREA_TUM_TOT",a_tum
write "Aria tumorala totala [micron^2] ="
write a_tum
pause

# Declarare si definire trasaturi de tip camp pentru masuratori: din FIELDFEAT – AREA_DIF_TOT...
# UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_DIF_TOT[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img8, care are rol de imagine-masca
RGnew 8,1

# Masurare trasaturi definite pentru toate regiunile din img8
MSmeasall "adk",0,2

# Scriere rezultate masuratori in fereastra Messages (fig. 4.5.9)
MSgetvalue "AREA_DIF_TOT",a_dif
write "Diferenta dintre aria tumorala totala si aria nucleara tumorala totala [micron^2] ="
write a_dif
pause

# Scriere rezultate masuratori in fereastra Messages (fig. 4.5.9)
# Definire arie nucleara tumorala totala
a_nctum = a_tum - a_dif
write "Aria nucleara tumorala totala [micron^2] ="
write a_nctum
pause

# Definire raport dintre aria nucleara tumorala totala si aria tumorala totala
rap_tum = a_nctum / a_tum
write "Raportul dintre aria nucleara tumorala totala si aria tumorala totala ="
write rap_tum
pause

# Declarare si definire trasaturi de tip camp pentru masuratori: din FIELDFEAT – AREA_STR_TOT...
# UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_STR_TOT[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img6, care are rol de imagine-masca
RGnew 6,1
```

Scanned with OKEN Scanner

```
# Masurare trasaturi definite pentru toate regiunile din img6
MSmeasall "adk",0,2

# Scriere rezultate masuratori in fereastra Messages (fig. 4.5.9)
MSgetvalue "AREA_STR_TOT",a_str
write "Aria stromala totala [micron^2] ="
write a_str
pause

# Declarare si definire trasaturi de tip camp pentru masuratori: din FIELDFEAT – AREA_DIFSTR_TOT...
# UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_DIFSTR_TOT[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img9, care are rol de imagine-masca
RGnew 9,1

# Masurare trasaturi definite pentru toate regiunile din img9
MSmeasall "adk",0,2

# Scriere rezultate masuratori in fereastra Messages (fig. 4.5.9)
MSgetvalue "AREA_DIFSTR_TOT",a_difstr
write "Diferenta dintre aria stromala totala si aria nucleara stromala totala [micron^2] ="
write a_difstr
pause

# Scriere in fereastra Messages (fig. 4.5.9)
# Definire arie nucleara stromala totala
a_ncstr = a_str - a_difstr
write "Aria nucleara stromala totala [micron^2] ="
write a_ncstr
pause

# Definire raport dintre aria nucleara stromala totala si aria stromala totala
rap_str = a_ncstr/ a_str
write "Raportul dintre aria nucleara stromala totala si aria stromala totala ="
write rap_str

beep 200,440
pause "FINAL ADK",0
stop
```

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **ADK**.

Scanned with OKEN Scanner

ANEXA II.5

- **MACROUL COREN**

```
# ********************************************************************************************
# NUME: COREN
# OBIECTIV: determinarea dimensiunilor (ariilor) ocupate de elementele componente ale unui corpuscul renal
# APELEAZA: fisierul imagine "cr.tif"
# RETURNEAZA: informatii numerice afisate in fereastra de mesaje
# BAZE DE DATE: coren.d00 (creare si utilizare)
# METODA: delimitarea regiunilor de interes prin segmentare RGB, segmentare tip crestere regiuni si operatii
# specifice imaginilor binare; masurarea automata prin utilizarea de trasaturi de camp
# ********************************************************************************************

# ---------------------------------------------------------------------------------------------------------------
# INITIALIZARE IMAGINI, PLAN GRAFIC SI FEREASTRA MESSAGES
# ---------------------------------------------------------------------------------------------------------------
# Initializare imagini
imgdelete "*"

# Incarcare imagine initiala (fig. 4.6.1)
imgload "c:\ks400\conf\images\photo\cr.tif",1

# Planul grafic devine transparent
Gclear 0

# Initializare fereastra Messages
showwindow "Messages", 1
write "@"
pause

# Setare parametri geometrici pentru masuratori: distance X = 68100/400 = 170.25 micro,
# distance Y = 47300/400 = 118.25 micro, scale X = 0.213, scale Y = 0.213
MSsetprop "SCALEX",0.213
MSsetprop "SCALEY",0.213
MSsetprop "UNIT","micrometer"

# ---------------------------------------------------------------------------------------------------------------
# PROCESARE IMAGINI
# ---------------------------------------------------------------------------------------------------------------
# Creare imagine binara noua in img10
# imgnew 10,803,558,1,"Grey"

# Imbunatatire contrast
brighten 1,2,0,75

# Afisare imagine
# imgdisplay 2

# Mod de lucru interactiv; se lanseaza Editorul grafic pentru desenarea conturului
showwindow "Editor",1

# Secventa de operatii: 1) apare mesajul; 2) se marcheaza creionul si culoarea; 3) se deseneaza contururile;
# 4) se deselecteaza editorul (creionul in afara figurii); 5) se minimizeaza img2; 6) se reactiveaza fereastra de
# mesaj in care se apasa OK. Editorul ramane deschis; desenul se realizeaza in planul grafic
MBok "Se traseaza contururi in jurul CR"
```

Scanned with OKEN Scanner

```
# Contururile din planul grafic se adauga la img2 (cu culoarea cu care au fost desenate) (fig. 4.6.2)
Gmerge 2,255

# Contururile din planul grafic se adauga la img10 (cu culoare alba) (fig. 4.6.3)
Gmerge 10,255

# Planul grafic devine transparent
Gclear 0

# Umplere gauri pe img10 cu depunere in img11
binfill 10,11

# Operatie booleana SAU-EXCLUSIV cu img10 si img11; rezultatul se depune in img12 (fig. 4.6.4)
binxor 10,11,12

# Segmentare RGB interactiva pe img2 pentru marcarea tuturor zonelor albe; rezultatul se depune in img3 (fig. 4.6.5)
dislevrgb 2,3,1,0,250,255,250,255,250,255,10,"RGB"

# Segmentare RGB interactiva pe img2 pentru marcarea nucleilor; rezultatul se depune in img4 (fig. 4.6.6)
dislevrgb 2,4,1,0,0,175,0,90,0,255,10,"RGB"

# Segmentare RGB interactiva pe img2 pentru marcarea matricei; rezultatul se depune in img5 (fig. 4.6.7)
dislevrgb 2,5,1,0,0,235,75,255,75,255,10,"RGB"

# Segmentare prin crestere regiuni pe img2 pentru marcarea spatiului Bowmann; rezultatul se depune in img6 (fig. 4.6.8)
reggrow 2,6,2,1,1,384,340,100,200,10

# Operatie booleana SI cu img12 si img3; rezultatul se depune in img13 (fig. 4.6.9)
binand 12,3,13

# Operatie booleana SI cu img12 si img6; rezultatul se depune in img7 (similar fig. 4.6.8)
binand 12,6,7

# Operatie aritmetica de scadere a img7 din img13; rezultatul se depune in img16 (fig. 4.6.10)
subtract 13,7,16,2

# Operatie booleana SI cu img12 si img4; rezultatul se depune in img14 (fig. 4.6.11)
binand 12,4,14

# Operatie booleana SI cu img12 si img5; rezultatul se depune in img15 (fig. 4.6.12)
binand 12,5,15

# ----------------------------------------------------------------------------------------------------
# ANALIZA CANTITATIVA
# ----------------------------------------------------------------------------------------------------
# Declarare nume pentru noua baza de date
DBnew "coren"

# Declarare si definire proprietati de camp pentru masuratori: din FIELDFEAT –
# AREA_TOT...UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_TOT[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img12, care are rol de imagine-masca
RGnew 12,1

# Masurare trasaturi de tip regiune declarate si definite anterior pe img12
MSmeasall "coren",0,2
```

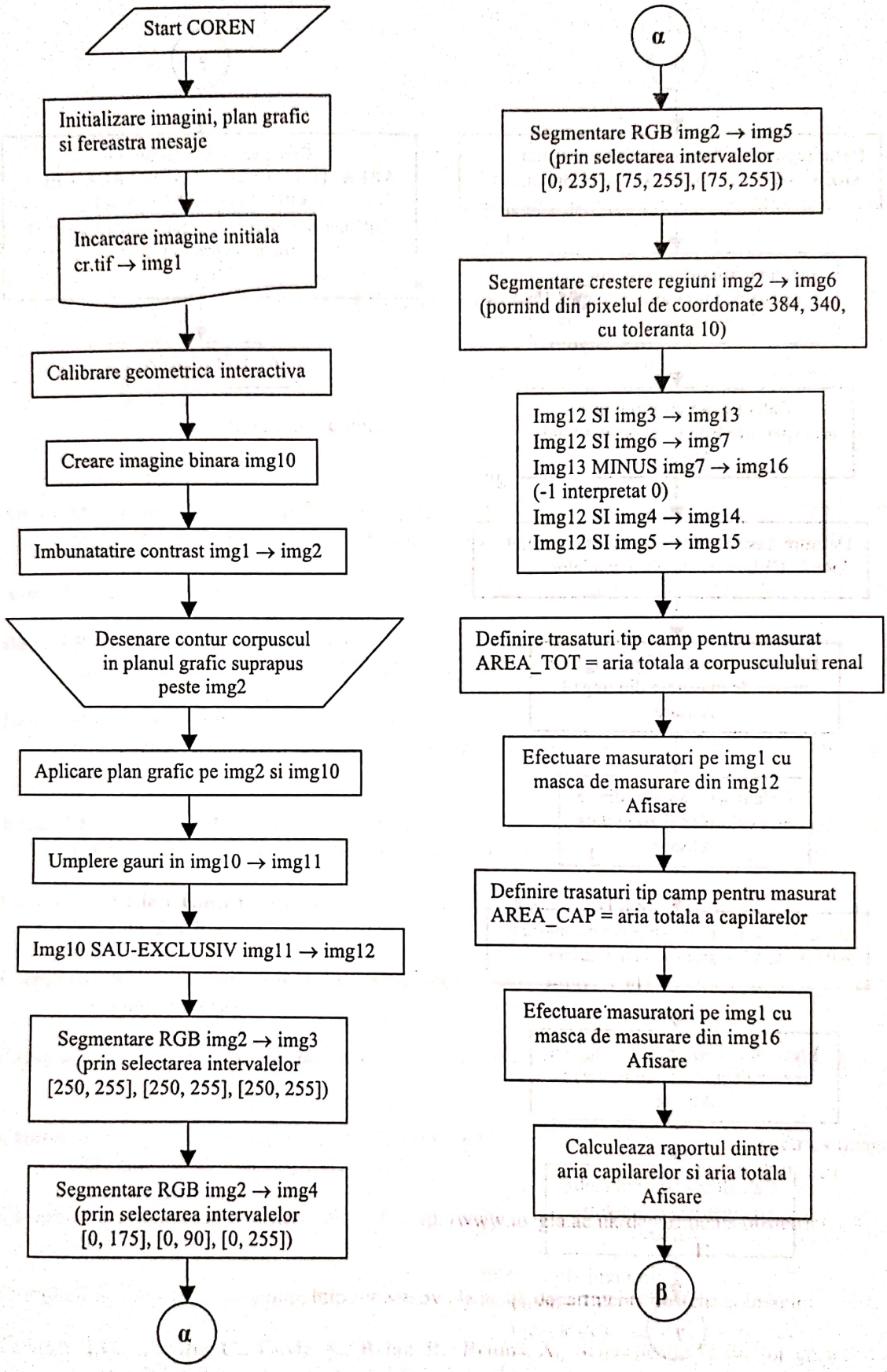
Scanned with OKEN Scanner

```
# Scriere rezultate masuratori in fereastra Messages
MSgetvalue "AREA_TOT",a_tot
write "Aria totala a corpusculului renal [micron^2] ="
write a_tot
pause

# Declarare si definire proprietati de camp pentru masuratori: din FIELDFEAT –
# AREA_CAP…UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_CAP[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img16, care are rol de imagine-masca
RGnew 16,1

# Masurare trasaturi de tip regiune declarate si definite anterior pe img16
MSmeasall "coren",0,2

# Scriere rezultate masuratori in fereastra Messages
MSgetvalue "AREA_CAP",a_cap
write "Aria totala a capilarelor [micron^2] ="
write a_cap
pause

# Definire raport dintre aria capilarelor si aria totala a corpusculului
rap_cap= a_cap / a_tot
write "Raportul dintre aria capilarelor si aria totala ="
write rap_cap
pause

# Declarare si definire proprietati de camp pentru masuratori: din FIELDFEAT –
# AREA_SB… UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_SB[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img7, care are rol de imagine-masca
RGnew 7,1

# Masurare trasaturi de tip regiune declarate si definite anterior pe img7
MSmeasall "coren",0,2

# Scriere rezultate masuratori in fereastra Messages
MSgetvalue "AREA_SB",a_sb
write "Aria totala a spatiului Bowman [micron^2] ="
write a_sb
pause

# Definire raport dintre aria spatiului Bowman si aria totala a corpusculului
rap_sb= a_sb / a_tot
write "Raportul dintre aria spatiului Bowman si aria totala ="
write rap_sb
pause

# Declarare si definire proprietati de camp pentru masuratori: din FIELDFEAT –
# AREA_CEL…UNIT^2...FLDAREA
MSsetprop "FIELDFEAT","AREA_CEL[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img14, care are rol de imagine-masca
RGnew 14,1

# Masurare trasaturi de tip regiune declarate si definite anterior pe img14
MSmeasall "coren",0,2
```

Scanned with OKEN Scanner

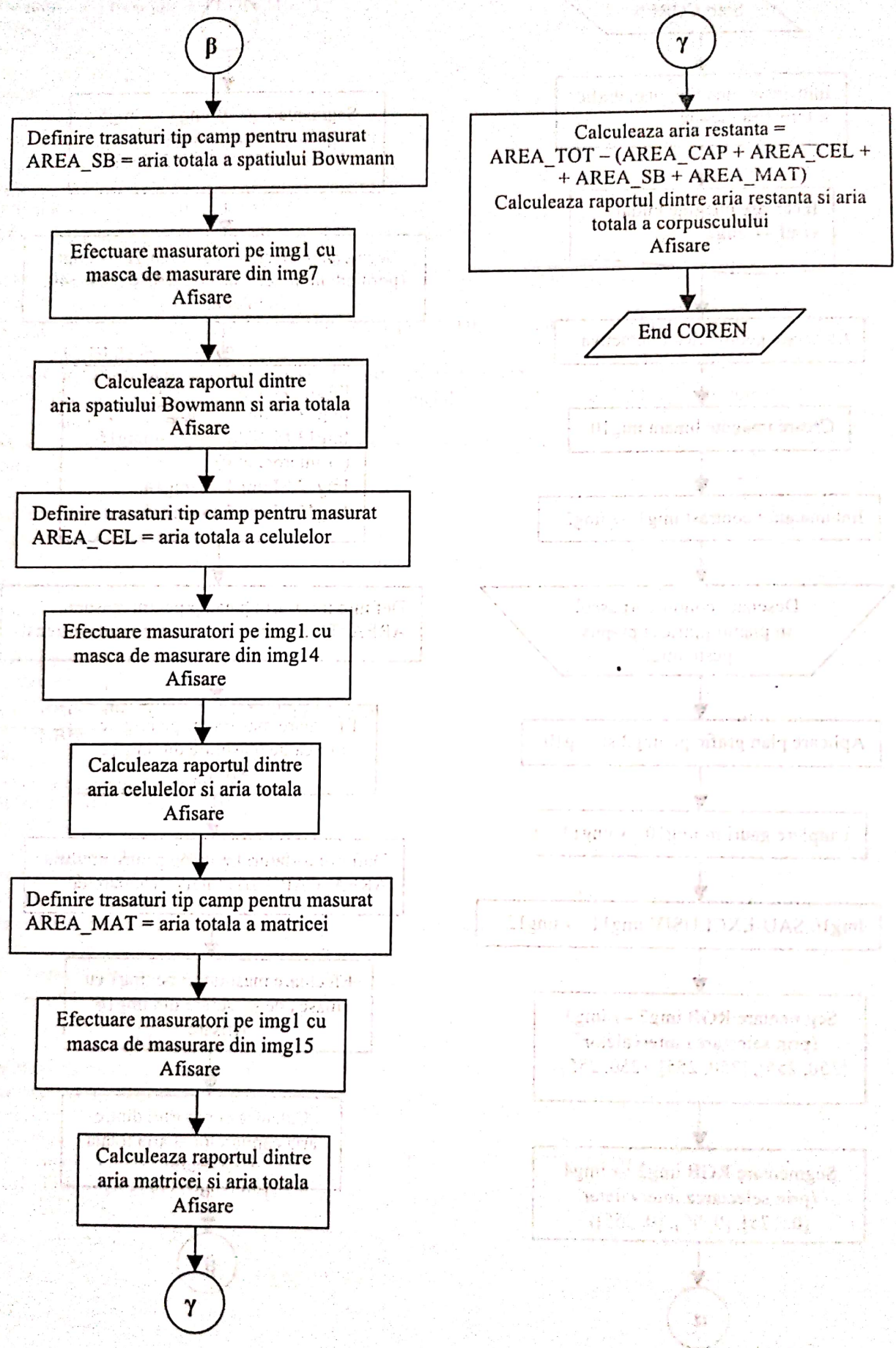
```
# Scriere rezultate masuratori in fereastra Messages
MSgetvalue "AREA_CEL",a_cel
write "Aria totala a celulelor [micron^2] ="
write a_cel
pause

# Definire raport dintre aria celulelor si aria totala a corpusculului
rap_cel= a_cel / a_tot
write "Raportul dintre aria celulelor si aria totala ="
write rap_cel
pause

# Declarare si definire proprietati de camp pentru masuratori: din FIELDFEAT –
# AREA_MAT...UNIT^2…FLDAREA
MSsetprop "FIELDFEAT","AREA_MAT[UNIT^2]=FLDAREA"

# Creare de lista de regiuni pe img15, care are rol de imagine-masca
RGnew 15,1

# Masurare trasaturi de tip regiune declarate si definite anterior pe img15
MSmeasall "coren",0,2

# Scriere rezultate masuratori in fereastra Messages
MSgetvalue "AREA_MAT",a_mat
write "Aria totala a matricei[micron^2] ="
write a_mat
pause

# Definire raport dintre aria matricei si aria totala a corpusculului
rap_mat = a_mat / a_tot
write "Raportul dintre aria matricei si aria totala ="
write rap_mat
pause

# Definire si calculare aria restanta
a_res=a_tot - (a_cap+a_cel+a_sb+a_mat)
write "Aria restanta [micron^2]="
write a_res

# Definire raport dintre aria restanta si aria totala a corpusculului
rap_res = a_res / a_tot
write "Raportul dintre aria restanta si aria totala ="
write rap_res
pause

beep 200,440
pause "FINAL COREN",0
stop
```

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **COREN**.

Scanned with OKEN Scanner



Schema logică pentru funcția utilizator **COREN** (continuare).

Scanned with OKEN Scanner

BIBLIOGRAFIE

Alvira M., Digital Imaging Theory, in: *Digital Imaging in Pathology* (Alvira M., Shireman P.K., Minarcik J.R.), ASCP National Meeting, Workshop # 9674, San Diego, California, 2000.

Avery J.K., *Oral Development and Histology*, 2nd ed., Thieme Medical Publishers, Inc., New York, 1994.

Baak J.P.A., The use and disuse of morphometry in the diagnosis of endometrial hyperplasia and carcinoma, *Path. Res. Pract.*, 179, 20-23, 1984.

Baak J.P.A. et al., Architectural and nuclear morphometrical features together are more important prognosticators in endometrial hyperplasias than nuclear morphometrical features alone, *J. Pathol.*, 154, 335-341, 1988.

Baak J.P.A., Artificial intelligence systems (expert systems) as diagnostic consultants for the cytologic and histologic diagnosis of cancer, *Cancer Res. Clin. Oncol.*, 114, 325-334, 1988.

Carl Zeiss Vision GmbH, *Imaging Components and Systems*, Halbergmoos, http://www.zeiss.de/Imaging, 2002.

Campbell S., *An Introduction to Stereology*, http://www.uv.gla.ac.uk/departments/obstetrics/Imaging/, 2002 (a).

Campbell S., *Introduction to the Digital Image*, http://www.uv.gla.ac.uk/departments/obstetrics/Imaging/, 2002 (b).

Campbell S., *Practical Imaging, part 1*, http://www.uv.gla.ac.uk/departments/obstetrics/Imaging/, 2002 (c).

Campbell S., *Practical Imaging, part 2*, http://www.uv.gla.ac.uk/departments/obstetrics/Imaging/, 2002 (d).

Campbell S., *Unbiased Sampling*, http://www.uv.gla.ac.uk/departments/obstetrics/Imaging/, 2002 (e).

Căruntu I.D., Cotuțiu C., Covic A., Balan R., Brumă A., Semnificația leziunilor crescentice în glomerulonefritele cronice, *Clinica*, II, 6, 7-11, 1997.

Căruntu I.D., *Morfogeneza sistemului stomatognat*, Ed. Apollonia, Iași, 2000 (a).

Scanned with OKEN Scanner

Căruntu I.D., *Album de histologie specială*, http://www.umfiasi.ro, 2000 (b).

Căruntu I.D., *Histologia sistemului stomatognat*, Ed. Apollonia, Iaşi, 2001 (a).

Căruntu I.D., *Oral Histology Album*, http://www.umfiasi.ro, 2001 (b).

Căruntu I.D., Highly effective techniques in computerized dental tissue morphometry, *J. Cell. Mol. Med.*, 6, 465-475, 2002.

Căruntu I.D., Microscopic morphometry – a modern approach, *Rev. Med. Chir. Soc. Med. Nat.*, 107, 1, 2003 (in press) (a).

Căruntu I.D., Analiza statistică şi de corelaţie în interpretarea rezultatelor măsurătorilor morfometrice, *Rev. Med. Chir. Soc. Med. Nat.*, 107, 2, 2003 (in press) (b).

Chmelar J., Sandvik K., A comparison of wet and dry autogenous grinding, *Mineral Processing on the Verge of the 21st Century*, 35-39, Balkema, Antalya, Turkey, 2000.

Connely D.P., Embedding expert systems in laboratory information systems, *Am. J. Clin. Pathol.*, 94, 507-514, 1990.

Collan Y. et al., *Morphometry of Kidney Biopsy: Number of Glomeruli and Accuracy of Morphometric Measurements. Stereology and Morphometry in Pathology*, Kuopio University Press, Kuopio, 1983.

Collan Y. et al., Scope and concepts of quantitative histopathology, *Acta Stereol.*, 11, 3-23, 1992.

Colgan T.J. et al., Predicting the outcome of endometrial hyperplasia by quantitative analysis of nuclear features using a linear discriminant function, *Int. J. Gynec. Pathol.*, 1, 347-352, 1983.

Davey D.F., *Morphometry*, http://www.physiol.usyd.edu.au/daved/teaching/emu, 2002.

Ehricke H., Grunert T., Buck T., Kolb R., Skalej M., Medical workstations for applied imaging and graphics research, *Computerized Medical Imaging and Graphics*, 18, 403-411, 1994.

Gonzales R.C., Woods R.E., *Digital Image Processing*, 2nd ed., Addison-Wesley Pub Co, 2002.

Gundersen H.J.G., Osterby R., Optimizing sampling efficiency of stereological studies in biology: or "Do more less well!", *J. Microscopy*, 121, 65, 1981.

Gundersen H.J.G. et al., Some new simple and efficient stereological methods and their use in pathological research and diagnosis, *APMIS*, 96, 379-394, 1988.

Hamilton P.W., Allen D.C., Morphometry in histopathology, *J. Pathol.*, 175, 369, 1995.

Howard C.V., Reed M.G., *Unbiased Stereology*, Bios Scientific Publishers, 1998.

Image Science Software GmbH, *IMAGIC-5*, Berlin, http://www.imagescience.de, 2002.

Institute of Biologic and Life Science – Image Analysis Centre, *Image Analysis Recent Work*, Glasgow, http://www.gla.ac.uk/ibls/ASU/imagan, 2002.

Iosifescu M., Moineagu C., Trebici V., Ursianu E., *Mică enciclopedie de statistică*, Ed. Ştiinţifică şi Enciclopedică, Bucureşti, 1985.

Scanned with OKEN Scanner

Jain K.A., *Fundamentals of Digital Image Processing*, Prentice-Hall International, Inc., Englewood Cliffs, New Jersey, 1989.

Jarchow R., *Image Analysis with KS 400*, http://www.multimedia.uni-greifswald.de, 2002.

Kindratenko V., *Development and Application of Image Analysis Techniques for Identification and Classification of Microscopic Particles* – Ph.D. Thesis, University of Antwerp, Belgium, 1997.

Kontron Elektronik GmbH, *Imaging System KS 400*, Rel. 2.0, User`s Guide, Munchen, 1996.

Kontron Elektronik GmbH, *Imaging System Zeiss KS400*, Munchen, http://www.kontron.com, 2002.

Macovei G.N., Varlam H., Antohe D. Şt., Analiza morfometrică a sistemului arterei cerebrale mijlocii la fătul de 13-15 săptămâni, *Rev. Med. Chir. Soc. Med. Nat.* (in press), 2002.

Malkusch W., Fundamentals and problem-solving sequences of quantitative image analysis, *Laboratory Focus*, 4, 2000.

Media Cybernetics, Inc., *OPTIMAS Image Analysis Software*, Silver Spring, Maryland, http://www.optimas.com, 2002.

Mihoc Ch., Micu N., *Elemente de teoria probabilităţii şi statistică*, Ed. Didactică şi Pedagogică, Bucureşti, 1971.

National Institute of Health – Research Services Branch, *NIH-Image*, Bethesda, Maryland, http://www.rsb.info.nih.gov, 2002.

Olinici C.D., Vaida M.F., *Metode de analiză cantitativă şi morfologică în biologie şi în medicină*, Ed. Tehnică, Bucureşti, 1997.

Pratt W.K., *Digital Image Processing*, 3rd ed., John Riley & Sons, 2001.

Rumşiski I.Z., *Prelucrarea matematică a datelor experimentale*, Ed. Tehnică, Bucureşti, 1974.

Russ J.C., *The Image Processing Handbook*, 3rd ed., CRC Press, Boca Raton, 1998.

Russ J.C., *The Image Processing and Measurement Cookbook*, http://www.reindeergraphics.com/tutorial, 2002.

Russ J.C., Dehoff R.T., *Practical Stereology*, 2nd ed., Plenum Press, 2001.

Seul M., O'Gorman L., Sammon M.J., *Practical Algorithms for Image Analysis: Description, Examples, and Code*, Cambridge University Press, 2000.

Shapiro L.G., Stockman G.C., *Computer Vision*, Prentice Hall, Upper Saddle River, New Jersey, 2001.

Slomianka L., *Blue Histology – A Brief Introduction*, http://www.teaching.anhb.uwa.edu.au/mb140/scope/stereology/stereint.htm, 2002.

Systems Planning and Analysis, Inc., *Stereologer*, http://www.stereologer.com, 2002.

Şabac I.Gh., Cocârlan P., Stănăşilă O., Topală A., *Matematici speciale*, vol. II, Ed. Didactică şi Pedagogică, Bucureşti, 1983.

Scanned with OKEN Scanner

Taşcă C., *Introducere în morfologia cantitativă cito-histologică*, Ed. Academiei Republicii Socialiste România, Bucureşti, 1976.

Ten Cate A.R., *Oral Histology. Development, Structure and Function*, 5th ed., Mosby, St. Louis, 1998.

The MathWorks, Inc., *Image Processing Toolbox for MATLAB*, Natick, Massachusetts, http://www.mathworks.com, 2002.

Tosi P., Cottier H., What`s new in quantitative pathology?, *Path. Res. Pract.*, 184, 652-655, 1989.

Tramontina V.A., Machado M.A.N., Kim S.H., Viterbo F., Nogueira-Filho G.R., Toledo S., Effect of fibrin glue (Tissucol) and fibrin glue prepaired from snake venom on palatal wound healing in rabbits, *IADR/AADR/CADR 80th General Session*, San Diego, 2002.

Turek J., *Introduction of Confocal Microscopy and Image Analysis*, Department of Basic Medical Sciences, School of Veterinary Medicine, Purdue University, 2000.

Văleanu I., Hâncu M., *Elemente de statistică generală*, Ed. Litera, Bucureşti, 1990.

Weibel E.R., Stereological principles for morphometry in electron microscopy, *Int. Rev. Cytol.*, 26, 235-302, 1969.

Weibel E.R., *Stereological Methods: Practical Methods for Biological Morphometry*, vol.1, Academic Press, London, 1979.

Weibel E.R., Measuring through the microscope: development and evolution of stereological methods, *J. Microscopy*, 155, 393-403, 1989.

Xu C., Lu M-G., Feng J-S., Guo Q-S., Wang Y-F., Germ cell apoptosis induced by Ureaplasma urealyticum infection, *Asian J. Androl.*, 3, 199, 2001.

Weisstein E., *Eric Weisstein's World of Mathematics*, http://www.mathworld.wolfram.com, 2002.

Scanned with OKEN Scanner

Cartea își propune să realizeze un studiu cuprinzător al tehnicilor digitale utilizate în procesarea informațiilor vizuale specifice preparatelor microscopice de histologie și histopatologie. Această abordare asistată de calculator este în principal orientată pe metodele moderne de morfometrie, care operează la nivel de pixel și asigură o acuratețe deosebită pentru analiza cantitativă. Prin modul în care este concepută, cu numeroase explicații, detalieri și exemple ilustrative, sfera cititorilor vizați cuprinde studenți cu aptitudini pentru activitatea de cercetare, studenți și tineri absolvenți preocupați de aplicabilitatea informaticii medicale, medici anatomo-patologi interesați de actualizarea cunoștințelor profesionale, prin prisma valorificării tehnicii de calcul.

ISBN 973-7906-34-9

Scanned with OKEN Scanner